

The Alberta Journal of Educational Research

Vol. IX, No. 4,

December, 1963



THE COMMITTEE ON EDUCATIONAL RESEARCH
Faculty of Education
University of Alberta

A STUDY OF THE EFFECTS OF THE "UNIFIED PHONICS METHOD" IN TEACHING READING IN GRADE I

A. FRED DEVERELL

*Associate Professor of Education
University of Saskatchewan*

Essentially, there are two, and only two ways by which young children can be introduced to reading: the "meaningful wholes" or sight-word approach, and the "component parts" or phonic method. All other methods are variations of a combination of these two. The currently most widely accepted "intrinsic method" is one which begins with "meaningful wholes", the body of sight-words which the child learns or memorizes in the initial stages—and subsequently teaches some elements of English phonics by a process of analysis of known sight-words. Both of these features continue throughout the first three years of the child's experience in learning to read, and, in many programs, into the fourth and fifth years.

Fundamental tenets of modern, orthodox reading instruction emphasized by most authorities include the following principles:

- (1) The child should learn a considerable body of words as sight-recognition words before any attempt at analysis is begun.
- (2) In his initial experiences in reading the child should read stories made up of previously learned sight words so that he will realize immediately that reading is a thoughtful process.
- (3) Any attempt to introduce phonics instruction, and to teach the child to sound out words, sound-by-sound and letter-by-letter, before he has grasped and had considerable experience with thoughtful story-type reading, will interfere with the development of meaningful reading, will create conscious word-calling and will produce a slow, inept type of reading.
- (4) Pictures and picture-reading are almost indispensable aids to the child in learning to read.
- (5) The child should not be taught to write and spell words until after he has experienced the words repeatedly in his reading.

The controversy which recently has centered on the relative merits of the two approaches to instruction mentioned above has arisen, at least in part, through dissatisfaction with results being achieved from the "intrinsic" method. It has led to widespread attention being drawn to a number of commercially available "phonetic" programs. One such program is that contained in the book, *The Writing Road to Reading*.¹

1. Romalda Bishop Spalding with Walter T. Spalding. *The Writing Road to Reading: A Modern Method of Phonics for Teaching Children to Read*, revised edition. New York: Whiteside and William Morrow, 1962. George J. McLeod, Ltd., Toronto.

The program was selected for special study for two principal reasons:

- (1) Its method is, in most of its characteristics, the exact anti-thesis of the currently orthodox "intrinsic" method, and it therefore provides an opportunity for an interesting study in contrasts.
- (2) It presents a program which stresses accurate speaking, spelling, writing, and reading as one integrated learning experience.

The program makes no use whatsoever of pictures; the child's attention is entirely directed, first to individual phonograms, and later to words and sentences. The child learns how to identify the sounds used in spoken English, and how to write them down as they are used in words. No words are read until the child has learned enough sounds to be able to identify the words by his own use of the letter and sound sequences. No books are read until the pupils have learned enough words to comprehend the meaning of the sentence. A total of seventy common phonograms, representing forty-five spoken sounds forms the core of this program. This program seemed to supply an interesting basis for a pilot study to assess the effects of the "Unified Phonics Method," as presented by the Spaldings in their book *The Writing Road to Reading*. Depending upon the outcome, a more extensive and formal study might then be developed.

The Problem

The stated problem involved specific definition of the word *effects* in the question, "What are the effects of the 'Unified Phonics Method' on pupils in first grade in learning to read?" The questions to which answers were sought through measurement by objective tests were:

- (1) How will the Unified Method affect the capacity of the children to recognize words?
- (2) How will the Unified Phonics Method affect the capacity of the children to read sentences and paragraphs meaningfully?
- (3) How will the Unified Phonics Method affect the capacity of the children to spell both familiar and unfamiliar words?

Other questions, the answers to which could be provided only by subjective evidence, were stated in the expectation that some useable evidence might be obtained. They were:

- (4) How will the Unified Phonics Method affect the development of ability to write, including the development of confidence in ability to reproduce language in writing?

- (5) How will the employment of the Unified Phonics Method affect the personality development of the children? Will the method and its materials prove to be drab and uninteresting? Will its failure to use interesting, colorful pictures have an adverse effect on the attitudes of the children toward reading and toward school life? Will the intensive training given in learning individual phonograms, before story reading can be introduced, create an inertia which will militate against successful learning? Under this program, will children become emotionally disturbed through the pressures of a difficult and demanding course which offers little in the way of rewards so much needed by beginners in a new task?

The Method of Study

Selection of Classes. The superintendent and the grade I teachers in the five public schools in the city of Yorkton, Saskatchewan, expressed interest in undertaking the pilot study. Two grade I teachers, Miss Baldwin of Fairview school, and Mrs. Coghill of Burke school, offered to teach the experimental program. The remaining teachers, Mrs. Brunsell and Mrs. Ingham of Simpson school; Mrs. Romanow and Miss Ellis of Angus Spice School, and Mrs. Adamson of Victoria school, were to be in charge of the control classes.

The Teaching Programs. A meeting was held with the two teachers of the experimental classes. It was agreed that they would proceed to teach by the Unified Phonics Method from the first day of school in September, and that they would, therefore, omit the formality of a reading readiness program, since the Unified Phonics Method makes no provision for such a program.

Since no materials are published for use with the Unified Phonics Method, other than a set of phonogram cards, it was agreed that the teachers of the experimental classes prepare their own teaching materials and make use of whatever book resources they could find. The first three books of the *Royal Road Readers*, by Daniels and Diack, sold in Toronto by Clarke, Irwin and Co., and published in Great Britain by Chatta and Windus, were chosen for use by both teachers as an aid in finding a common ground for initial book reading. However, these teachers insisted upon avoiding the employment of any of the standard pre-primers and primers in common use in the schools because of the opposing nature of philosophy of teaching in these books. Finally, it was agreed that the pupils in the experimental classes would, before the year-end, read the primer and first reader of the prescribed readers for Saskatchewan as supplementary reading. This was done to give assurance to the Director

of Curriculum of the province that the children in the experimental classes would get at least as effective preparation as is given to all other pupils in the province.

The control classes were to be instructed in the regular, prescribed program as taught in the *Curriculum Foundation Series* of readers, beginning with a readiness program and proceeding in the normal way through the pre-primers, primer, first reader, and supplementary readers, and employing the teaching program as set out in the guide books and work-books as well as whatever modifications any of these teachers had developed in using this program.

Testing. A pre-test: post-test plan was adopted. Pre-tests selected were the *Watson Reading-Readiness Test*,² to be administered in all classes at about the end of September, and the *Dominion Group Test of Learning Capacity (Primary)*,³ to be administered in all classes at the end of October or early in November. Late in February, re-measuring of all classes was to be done by the use of two tests. Reading abilities would be measured by the use of the *Gates Primary Reading Tests*, Types PWR, PSR and PPR. The investigator undertook to prepare a test which was called the *Word and Sentence Test*, to be used for measuring capacity in spelling and handwriting. A sample of this test has been provided in Appendix A. No tests were selected for the measurement of personality and adjustment factors.

Limitations of the Experiment

The limitations of an experiment in which neither the teachers employed in experimental and control groups, nor the classes, have been randomized are obvious. One of the two teachers who volunteered to teach the experimental program did so because she favored the program; the other because she felt strongly opposed to it. This may, to some degree, balance the teacher factor, but it does not eliminate the possibility that both experimental teachers were more than usually competent. The ultimate results show that there was certainly considerable competence in the teachers of the control classes.

Presentation of Results

Only pupils who were present for all tests in the pre-test and post-test series of tests were used in the final analysis of test data, as presented here. The result is that, although there were originally fifty-two pupils in the experimental group and one hundred and twenty-nine pupils in the control group, the N's ultimately used were

2. G. Milton Watson. *The Watson Reading-Readiness Test*. Agincourt, Ontario: The Book Society of Canada, 1959.

3. *The Dominion Tests: Group Test of Learning Capacity (Primary)*, for Kindergarten, Kindergarten-Primary, and Grade I, Form A, revised edition. Toronto: Department of Educational Research, Ontario College of Education, 1956.

forty-one and one hundred seventeen, respectively. Frequency tables of test scores are shown in Appendix B.

Equivalence of groups. The results of the *Watson Reading Readiness Test* and the *Dominion Test of Learning Capacity* were used to test the equivalence of the experimental and control groups. Since some negative skewness was observable in the distribution of scores on both of these tests, the median test of significance of difference was used. The data are presented in Tables I and II.

In both cases the values obtained for χ^2 are smaller than those required to establish a significant difference between the two groups at either the .01 or the .05 level of significance. The null hypothesis that there is no significant difference between the two groups in scores in the two pre-tests is therefore sustained. These tests indicate that the two groups came from the same population.

TABLE I
CHI SQUARE TEST OF SIGNIFICANCE OF DIFFERENCE
BETWEEN THE MEDIAN OF THE EXPERIMENTAL AND
CONTROL GROUPS IN THE WATSON READING
READINESS TEST

	Experimental	Control	Total
Above Median	19	60	79
Not above median	22	57	79
Total	41	117	158

$$\chi^2 = \frac{158 \left[(19)(57) - (60)(22) - \frac{158^2}{2} \right]}{(79)(79)(41)(117)} = 0.527$$

- (1) Ho: The two groups have the same median.
- (2) Level of significance: .01 or .05.
- (3) Region of rejection: $\chi^2_{01} \geq 6.635$

(1 d.f.)

$\chi^2_{05} \geq 3.841$
- (4) Observation: $\chi^2 = 0.527$.
- (5) Conclusion: The hypothesis is sustained. There is no significant difference between the medians of the two groups.

TABLE II

CHI SQUARE TEST OF SIGNIFICANCE OF DIFFERENCE
BETWEEN THE MEDIAN OF THE EXPERIMENTAL AND
CONTROL GROUPS IN THE DOMINION TEST OF
LEARNING CAPACITY

	Experimental Group	Control Group	Total
Above Median	25	53	78
Not above Median	16	64	80
Total	41	117	158

$$\chi^2 = \frac{158 \left[(25)(64) - (53)(16) - \frac{158^2}{2} \right]}{(78)(80)(41)(117)} = 2.391$$

- (1) Ho: The two populations have the same median.
- (2) Level of significance: .01 or .05.
- (3) Region of rejection: $\chi^2_{01} \geq 6.635$

(1 d.f.)
 $\chi^2_{05} \geq 3.841$
- (4) Observation: $\chi^2 = 2.391$
- (5) Conclusion: The hypothesis is sustained. There is no significant difference between the medians of the two groups.

Post-test measure of achievement. After six months of instruction time had elapsed the achievement of both groups was measured by administering the *Gates Primary Reading Test*, Types PWR, PSR, and PPR, and the *Word and Sentence Writing Test* prepared by the investigator (Appendix A).

The Chi square test of significance of difference between the medians in the performance of the two groups, based on the null hypothesis of no difference between the medians, was employed. The results are reported in Table III.

In all four measures the hypothesis that there was no difference between the medians of the two groups was rejected. In sentence reading and total Gates raw score the region of rejection was at the .001 level of significance; in word recognition and paragraph reading rejection occurred at the .01 level of significance. The experimental group, therefore, appears to have performed in this test at a level which is significantly superior to that of the control group.

TABLE III
CHI SQUARE TESTS OF SIGNIFICANCE OF DIFFERENCE
BETWEEN THE MEDIANS OF THE EXPERIMENTAL AND
CONTROL GROUPS IN SUB-TEST RAW SCORES AND
TOTAL RAW SCORES OF THE GATES PRIMARY READING
TEST: NULL HYPOTHESIS

Test	Obtained χ^2 Value*	Region of Rejection of the Null Hypothesis at Indicated Level of Sig- nificance.	Conclusion
(1) Word Recognition	10.672	$\chi^2_{.01} = 6.634$ (1 d.f.)	Hypothesis rejected†
(2) Sentence Reading	12.502	$\chi^2_{.001} = 10.827$ (1 d.f.)	Hypothesis rejected††
(3) Paragraph Reading	6.944	$\chi^2_{.01} = 6.634$ (1 d.f.)	Hypothesis rejected†
(4) Total Gates Raw Score	12.502	$\chi^2_{.001} = 10.827$ (1 d.f.)	Hypothesis rejected††

$$*\chi^2 = \frac{N \left[(a)(d) - (b)(c) - \frac{N}{2} \right]^2}{(k)(l)(m)(n)}$$

†The difference is significant at the .01 level and favors the Experimental Group.
††The difference is significant at the .001 level and favors the Experimental Group.

Marked negative skewness was observed in the distribution of scores in the *Word and Sentence Writing Test*, and the Chi Square test was employed here also to test the significance of difference between the medians of the two groups. Table IV shows that the performance of the experimental group was superior to that of the control group in the *Word and Sentence Writing Test* by a difference which was very highly significant.

Analysis of Errors in the Word and Sentence Test. A more informative type of analysis of the differences in the performance of the two groups might be achieved by ascertaining the types of errors made by the pupils and determining whether or not there was any significant difference in the performance of the two groups in respect to these errors. It was found that all errors could be classified into six categories: reversals, rotations, consonant confusions, vowel confusions, omissions and additions. A *reversal* is a transposition of the order of letters as in writing on for no; frist for first. A

TABLE IV

CHI SQUARE TEST OF SIGNIFICANCE OF DIFFERENCE
BETWEEN THE MEDIANS OF THE EXPERIMENTAL AND
CONTROL GROUPS IN THE WORD AND SENTENCE
WRITING TEST

	Experimental Group	Control Group	Total
Above Median	31	47	78
Not above Median	10	70	80
Total	41	117	158

- (1) Ho: The two groups have the same median.
- (2) Level of significance: .01 or .001.
- (3) Region of rejection: $\chi^2_{01} \geq 6.635$
(1 d.f.)
 $\chi^2_{001} \geq 10.827$
- (4) Observation:
- $$\chi^2 = \frac{158 \left[(31)(70) - (47)(10) - \frac{158^2}{2} \right]^2}{(78)(80)(117)(41)} = 13.870$$
- (5) Conclusion: The hypothesis is rejected. There is a significant difference between the medians of the two groups, at the .001 level of significance. The difference favors the experimental group.

rotation is a turning of a letter on its vertical or horizontal axis as in writing b for d; u for n; d for g. Consonant confusions include errors like writing w for wh, or th for t, or sl for sh; or r for l. Vowel confusions include many possibilities such as writing a for e or e for a; writing i for e or e for i; writing u for w or w for u; writing a for ai or for a-e; writing a-e for ai, etc. Omissions are chiefly the omission of a consonant letter; omission of an e needed at the end of a word to make the medial vowel long classes as a vowel confusion. Likewise, additions usually involve the addition of a consonant, or of a syllable involving at least a vowel and a consonant. All of the errors made could be readily classed as one of these six types. The only other possible type of error was a complete *refusal* to write the word, an error which was committed so rarely as to be unworthy of assessment.

The significance of difference between the incidence of the above six types of error in the experimental and control groups was tested

by means of the 2×2 contingency table and the Chi Square test, using one degree of freedom. In the case of each error the number of pupils making *no errors* was tested against the number making *one or more errors* of the particular type. The results are summarized in Table V.

In Table V it is seen that the difference in performance of the two groups with respect to *reversal errors* was significant at the .01 level. This difference favored the experimental group. With respect to *rotations*, *consonant confusions*, and *vowel confusions*, the per-

TABLE V
CHI SQUARE TESTS OF SIGNIFICANCE OF DIFFERENCE
BETWEEN EXPERIMENTAL AND CONTROL GROUPS IN
COMMITTING ERRORS OF SIX TYPES IN THE WORD AND
SENTENCE WRITING TEST, USING THE HYPOTHESIS OF
NO DIFFERENCE*

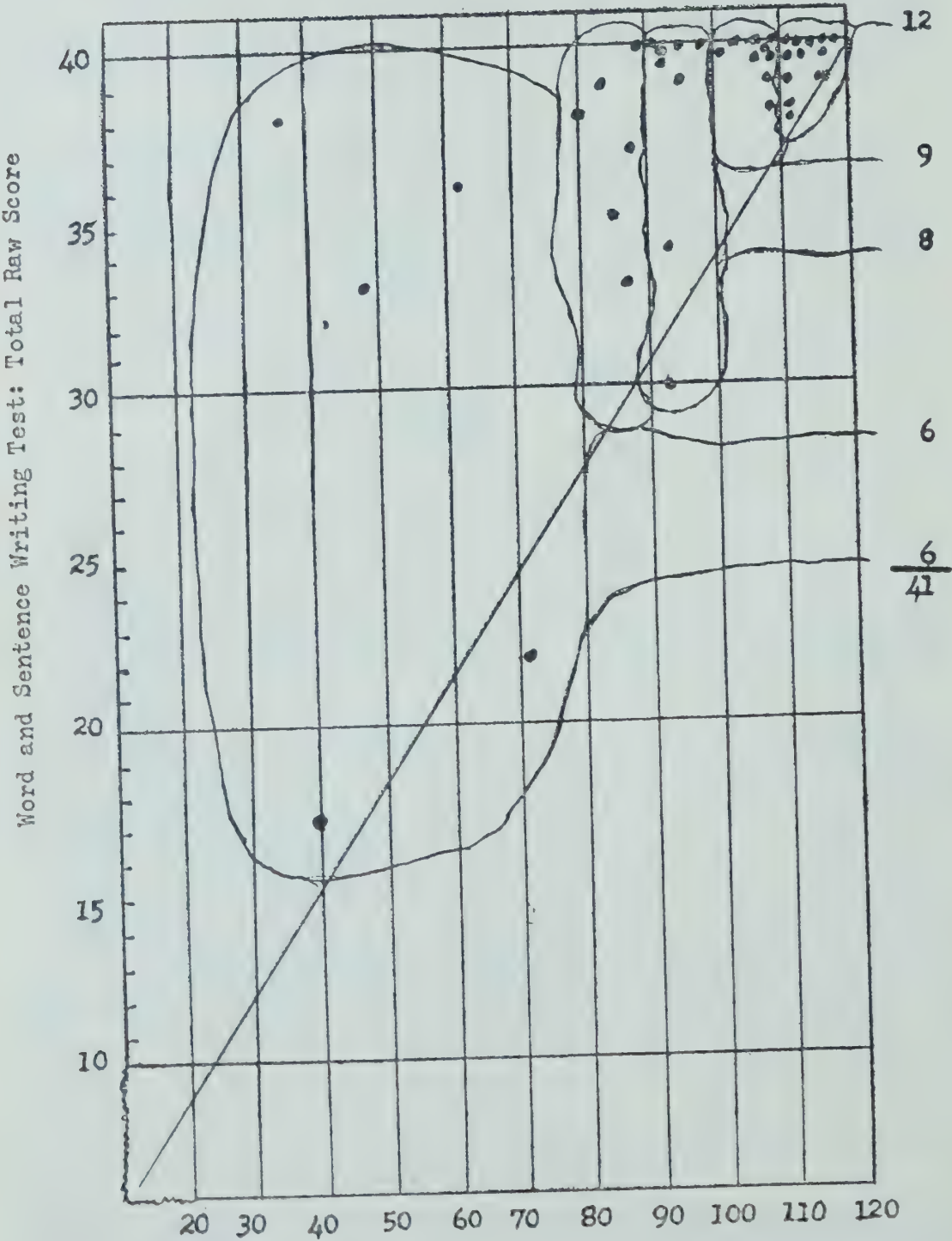
Type of Error	Obtained χ^2 Value	Region of Rejection of the Null Hypothesis at Indicated Level of Significance	Conclusion
(1) Reversals	8.275	$\chi^2_{.01} = 6.635$ (1 d.f.)	Hypothesis rejected†
(2) Rotations	4.458	$\chi^2_{.05} = 3.841$ (1 d.f.)	Hypothesis rejected†
(3) Consonant Confusions	3.912	$\chi^2_{.05} = 3.841$ (1 d.f.)	Hypothesis rejected†
(4) Vowel Confusions	5.823	$\chi^2_{.05} = 3.841$ (1 d.f.)	Hypothesis rejected†
(5) Omissions	3.266	$\chi^2_{.05} = 3.841$ (1 d.f.)	Hypothesis sustained
(6) Additions	1.126	$\chi^2_{.05} = 3.841$ (1 d.f.)	Hypothesis sustained

*(1) χ^2 was calculated by means of 2×2 contingency tables in which the number of pupils making no errors were tested against the number of pupils making one or more errors.
(2)

$$\chi^2 = \frac{N \left[(a)(d) - (b)(c) - \frac{N}{2} \right]^2}{(k)(l)(m)(n)}$$

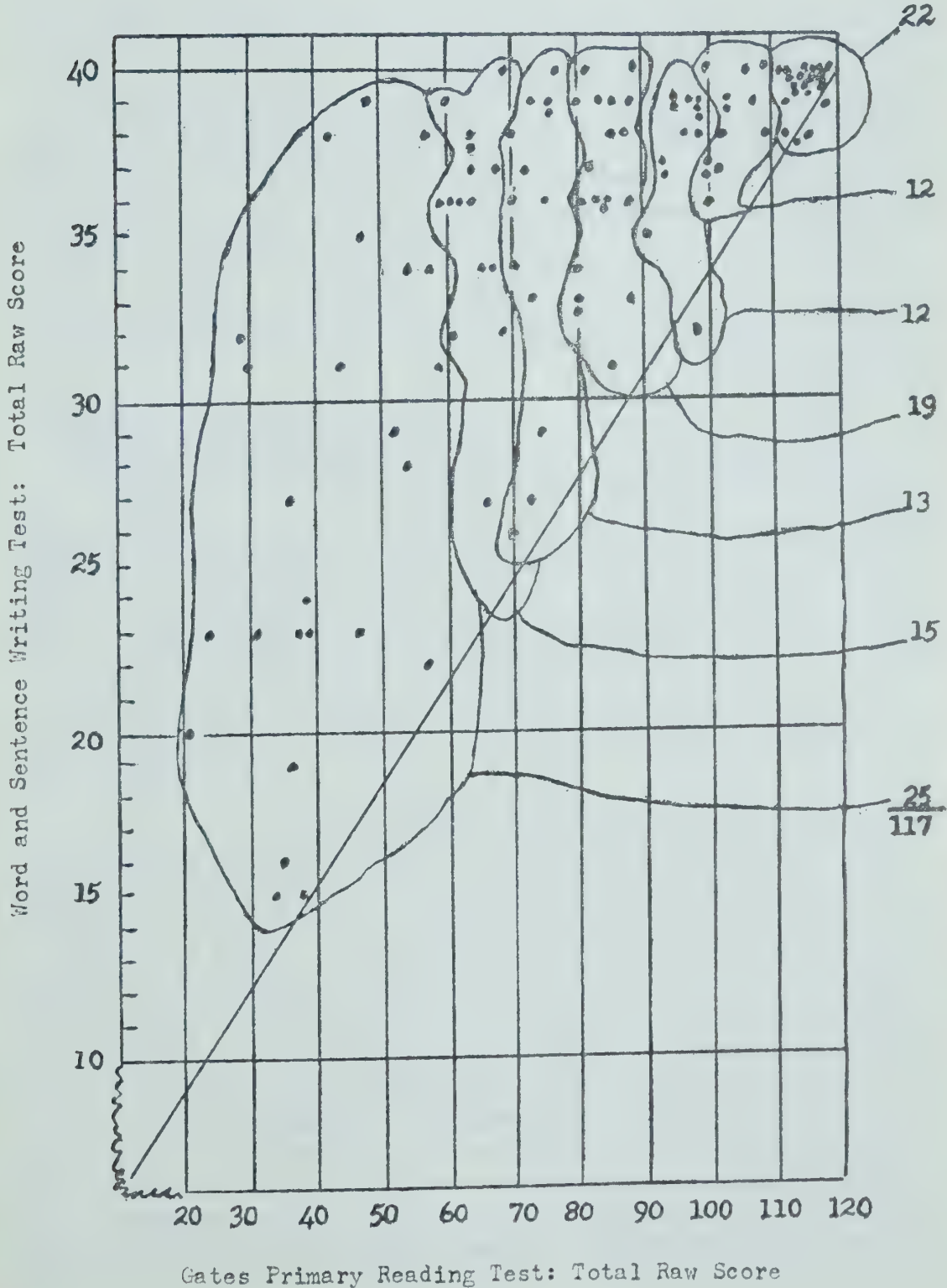
†In all instances in which the null hypothesis was rejected the difference favored the experimental group. The experimental group made significantly fewer errors in the first four categories of error.

TABLE VI. Scattergram, Showing Distribution of 41 Pupils in the Experimental Group on the Total Score of the Gates Primary Reading Test and the Word and Sentence Writing Test



Gates Primary Reading Test: Total Raw Score

TABLE VII. Scattergram, Showing Distribution of 117 Pupils in the Control Group on the Total Score of the Gates Primary Reading Test and the Word Sentence Writing Test



formance of the experimental group was superior to that of the control group, the difference being significant at the .05 level. There was no significant difference with respect to *omissions* and *additions*.

Scattergrams. Scattergrams were prepared to study the relationship between scores in the *Gates Primary Reading Test* (Total raw score) and the *Word and Sentence Writing Test*. These are presented in Tables VI and VII. Except for the cases of two pupils in the experimental group and four pupils in the control group, the lower right half of each scattergram is empty. No pupils who scored low in spelling (the *Word and Sentence Writing Test* is a test in spelling) scored high in reading. The converse of this statement is not true; there are pupils who scored low in reading but high in spelling. From these limited data it is possible to make the tentative generalization that *ability to spell is a necessary, but insufficient condition for good reading*.

Assuming curvilinearity in both distributions, Eta coefficients of correlation were calculated. For the experimental group, $N_{yx} = .83$; for the control group, .54. These correlations were not tested for significance.

Other criteria for evaluating the effects of the Unified Phonics Method. Initially (page 3 and 4) it was stated that an attempt would be made to assess the effects of the Unified Phonics Method on the ability of the pupils to write and to express themselves in writing. It was also stated that some attempt would be made to estimate the effects of this method on the personality development of the children, on the development of interest and in enjoyment of reading and school life. Aside from rather casual observation, no firm attempt was made to evaluate these criteria.

The writing done by the pupils in taking the *Word and Sentence Writing Test* provided a writing sample for each pupil. In general it may be said that pupils in the experimental group produced papers in which there was more frequent excellence in the quality of the manuscript writing than was true of the control group pupils. There was less evidence among pupils in the experimental group of the tendency to press excessively on the pencil, to make corrections and erasures, and to exhibit difficulty in writing. There was also in the experimental group a greater degree of correctness in letter formation and spacing and of uniformity of slant than was evident in the papers from the control group. A very large proportion of the papers written by the experimental group pupils were models of excellence. The boys tended to write as well as the girls. By contrast, in the control group there was much excellence but not as high a pro-

portion of it, and more pupils exhibited lack of ease and fluency in their handwriting.

Samples of creative free writing were taken from all pupils in this experiment in June. At the time of writing this report they had not been analysed.

Summary

An experiment was conducted, as a pilot study, in the schools of Yorkton, Saskatchewan, in the school year 1962-63 to assess the effects of the Unified Phonics Method as expounded in *The Writing Road to Reading*, by Spalding and Spalding. Two teachers volunteered to teach the experimental program; the five remaining grade I teachers taught by their regular procedures.

The equivalence of the two groups of pupils was tested by means of the *Watson Reading Readiness Test* and the *Dominion Test of Learning Capacity*. In both tests it was established that the pupils represented the same population.

Achievement in reading was tested after six months of instruction by use of the *Gates Primary Reading Test*. The level of the experimental group exceeded that of the control group in all measures. The differences favoring the experimental group were significant at the .001 level of confidence in Total Score and Sentence Reading, and almost at that level in Word Recognition. The difference was significant at the .01 level in Paragraph Reading.

The Gates Primary Reading Test was not adequate as a measuring instrument; too many pupils earned the maximum scores in the various sub-tests.

Capacity of the pupils in spelling was measured by the *Word and Sentence Writing Test* prepared by the writer. The results for the experimental group were superior to those of the control group, the difference being significant at the .001 level of significance.

The errors made by the pupils in spelling were analyzed. The experimental group made significantly fewer errors of the Reversal type (.01 level of significance) and in Rotations, Consonant Confusions and Vowel Confusions (.05 level of significance). There was no significant difference in Omissions and Additions.

The scattergram analysis of relationship between reading and spelling scores gave strong evidence that spelling ability is an important concomitant of reading ability. No poor spellers were good readers, although some good spellers were poor readers. The level of the correlation in relation to the experimental group was .83; in the control group .54. Evidently a method in teaching reading which

stresses writing and spelling produces a higher level of consistency of performance in all three abilities.

Pupils in the experimental group, as a whole, were better writers and wrote with more fluency and with less tension than pupils in the control group. This is not to suggest that there were not many good writers in the control group and some less fluent ones in the experimental group.

Conclusions

1. In this experiment, a method in teaching reading in grade I which had no readiness program, which employed no readers and no pictures, and which stressed articulated phonics, writing and spelling as integral parts of learning to read, produced results significantly superior to those of the regular teaching procedures.
2. This experiment involved too few teachers and pupils to permit broad generalizations. Its results justify a large-scale, longer-term experiment.
3. If, ultimately, the results of this study are confirmed by repeated experiments, there will then be confirmation of the belief of the writer that present instruction procedures are faulty in separating the learning of writing and spelling from the learning of reading, in the wrong use of pictures, and in the too limited scope of the reading matter given to children as practice material.
4. The claim made by the authors in *The Writing Road to Reading* that *all* pupils learn to read by the Unified Phonics Method was not sustained in this experiment. There are at least four pupils whose achievement is included in the analyses of results presented in this report whose level of achievement cannot be regarded as one of mastery. In addition to this there were two pupils whose results were excluded by reason of absence for one or more of the tests, and who made very little, if any progress during the six months of instruction. The cases of these two pupils deserve most careful study, as both give evidence of severe interference with the normal processes of perception and cognition.
5. In any repetition of this experiment specific provision should be made for obtaining an index of interest in reading and the desire to read among experimental and control group pupils. Provision should also be made for recording of teacher reactions and parent observations concerning the welfare of the pupils under the two systems of instruction.

APPENDIX A

COLLEGE OF EDUCATION
UNIVERSITY OF SASKATCHEWAN



Word Writing and Sentence Writing Test for Grade I
Prepared Especially for Use in the Yorkton, Saskatchewan, Schools
In February or March, 1963, by A. F. Deverell



INSTRUCTIONS

This test involves a total of 40 written words and is designed to be given in two parts on two separate days. Information concerning the validity of the test items is provided on a separate attached sheet. In brackets after each word the level at which the word is introduced in the Basic Readers is indicated. The test is designed as a supplement to the standardized reading tests in regular or periodic use. It is intended to test the capacity of the children to reproduce, in print writing, the words and the word elements which they are learning in their reading.

How to Give the Test:

- (1) Tell the children that you are going to give them some paper on which you will ask them to write some easy words.
- (2) Pass out foolscap sheets and show the children how to fold the sheet "in twos" to make two long columns from the top to the bottom of the page. (If you wish you may have them use ruler and pencil to make a vertical line down the centre of the sheet along the fold.)
- (3) Give the instructions you ordinarily do regarding the spacing of their writing. (Two-space-high writing is assumed.)
- (4) Have each pupil write his name (both given and surname) at the top of the page.
- (5) Tell the children that you will say each word they are to write, then use it in a sentence, and then that you will say the word again. Say to them, "Listen carefully for the word. I want you to write exactly the word you hear me say. Write it exactly like it sounds when I say it." (If you need to repeat the word again once, or several times for any child to be sure he has heard it, then do so. This is not a hearing test. It is a test of how the children can write these words.)

WORD AND SENTENCE WRITING TEST, GRADE I
PART I

City Date of Test
(See directions for scoring on page 2 of the "Instructions".)

SCHOOL TEACHER NO. IN CLASS

		Dictation Words	Number Having Each Word Right	Number Having Each Word Wrong	Number Who Omitted Each Word
1. BABY	(pp2)	I have a baby sister.			
2. DOG	(p)	My dog plays with me.			
3. CHILDREN	(p)	The children go to school.			
4. FUN	(p)	We have fun at school.			
5. HELP	(pp2)	I like to help Mother			
6. IN	(pp3)	Put the pencil in the box			
7. LIKE	(p)	We like to write our words.			
8. JANE	(pp1)	Jane likes to write her words.			

Now, put No. 9 at the top of the next column on your sheet.

9. MUST	(p)	Soon we must go home for lunch.			
10. NOT	(pp3)	Tom will not go with her.			
11. PETS	(p)	We like to feed our pets.			
12. QUACK	(p)	The duck said. "Quack."			
13. RED	(pp2)	My pencil is red.			
14. SAW	(p)	I saw Tim in the store.			

Dictation Sentence (Write this sentence across the bottom of your page.)							
I (pp1)	am (p)	playing (l-p132)	with (p)	my (pp2)	toys (p)		
					15. I		
					16. AM		
					17. PLAYING		
					18. WITH		
					19. MY		
					20. TOYS		
Totals:							
Percent of words which are right, wrong, and omitted:							

NOTE: This sheet has been arranged so that, after scoring the individual papers, a record can be made of the number of times each word was written correctly, incorrectly, or omitted. The horizontal totals for each word should be the same as the number of pupils in the class. If you have any pupil who can write no words at all, exclude him (or her) from your class total, and from the count of words omitted, but include his (or her) paper with the others when sending in the test results. Mark the paper in red pencil—NOT SCORED.

WORD AND SENTENCE WRITING TEST, GRADE I
PART II

SCHOOL

TEACHER

NO. IN CLASS

(Directions for administering and scoring are the same as for Part I.)

City		Date of Test		
Dictation Words				
		Number Having Each Word Right	Number Having Each Word Wrong	Number Who Omitted Each Word
1. VERY	(1-21) Tom was a very big boy.			VERY
2. WAS	(p) Tom was in our school.			WAS
3. THREE	(pp3) The cat had three kittens.			THREE
4. WHITE	(p) One kitten was white.			WHITE
5. YES	(p) The boy said, "Yes, I will help."			YES
6. BARN	(p) The horse is in the barn			BARN
7. BOAT	(pp2) We ride in our boat.			BOAT
8. ALL	(p) I want all my money.			ALL

Now, put No. 9 at the top of the next column on your sheet.

9. SHE	(p) Mary said that she could jump.	SHE		
10. COW	(p) Our cow gives us milk.	COW		
11. EAT	(p) Cows like to eat grass.	EAT		
12. MEW	(p) The kitten said, "Mew."	MEW		
13. TAIL	(p) Our cat has a long tail.	TAIL		
14. PLAY	(pp2) We want to go out to play.	PLAY		
15. WORK	(pp2) My daddy likes to work.	WORK		

Now, I am going to write a sentence on the board. Watch me while I write it. Then I want you to write it across the bottom of your sheet.				
We (pp3)	went (p)	to (pp3)	the (pp2)	house. (pp3)
16. WE				
17. WENT				
18. TO				
19. THE				
20. HOUSE				
Totals:				
Per cent of words tried which were right, wrong, and omitted:				

- (6) As you dictate the word, give the number of each word and have the children number their words down the page. Including the words in sentences, there are 20 words in each part of the test. Have them write the first 8 words on the left side of the sheet, and the rest of the column words on the right side.
- (7) When you come to the sentence dictation, tell them to write the sentence right across the page in the spaces near the bottom. Read the *whole sentence* slowly first, and then repeat it, word-by word, for clarity. Re-read and let the pupils check after they have completed writing it.
- (8) *In Part II the sentence is to be put on the board by the teachers, and the pupils are to be instructed to copy the sentence at the bottom of the page.*

The teacher should give all the friendly encouragement she can to the pupils, and help those who have difficulty getting the words in column to understand what to do. Pretend you see no errors, and encourage all to write every word. *Give no help in how to write a word or letter.*

Directions for Scoring and Tabulating:

- (1) Mark each word on each test paper right (✓) or wrong (×), or omitted (O). Ignore capitals. Accept the right symbol, whether upper or lower case.
- (2) Record on the test sheet the number who took the test. If you have any pupil who could write no words whatever, exclude that pupil from your total, but include his paper with the set of written test papers so that any marks which he made on the paper can be seen. If you have any pupil who fails to write all or most of them, but who attempts some or many of them, include such pupil and record on the summary of errors sheet what he (or she) actually put down for each word.
- (3) Complete the record on the test paper.
- (4) Complete the Diagnostic Record Sheets for your class.
- (5) Place the test papers, together with these records, in an envelope and return as directed.

Your cooperation in administering and scoring this test is much appreciated. You will be fully advised of the outcomes of this test and other tests in this series.

WORD AND SENTENCE TEST, GRADE 1

Notes on the Content Validity of the Test

1. Every effort has been made to construct a test which is valid (i.e. fair or justifiable) for grade 1 pupils in mid-year. It would be comparatively simple to construct a test which is highly valid for the children taught by the basic reader method or to construct one which is valid for children taught by the "Writing Road to Reading" method. To make a single test which has equal validity for both groups of children is a much greater challenge. The following information will help to judge the content validity of this test for both groups of children.
2. The words chosen for inclusion in the test are all words which were introduced in the pre-primers and the primer of the Basic Reader, with one exception. This one word is introduced in the first reader. This means that, with one exception, the children in the control group will have encountered all of the words many times in their reading and in their work-book exercises. Pupils in experimental classes may, or may not have encountered these particular words, but will be familiar with the phonic elements in the words.
3. The words were so selected as to include all of the single consonant symbols in English except "c" used with the "s" sound, and the consonants "x" and "z". These three symbols were omitted because they do not appear in the vocabulary of the pre-primers and primer. Pupils in the experimental group will have encountered these symbols, even though they are not tested.
4. The words were so chosen as to present both the short vowel sounds (when closed by a consonant) and the long vowel sound (when, open, or not closed, by a consonant), for the vowels "a", "e", and "i". Only the short sound (when closed) of the vowels "o" and "u" have been included. The vowel made long by silent "e" at the end of the word has been included only for the vowels "a" and "i".
5. The vowel digraphs included are ai, ay, ee, oa, and oy. The vowel-consonant digraphs included are aw, wa, ar, ew, er, ow, ou, and or. The consonant digraphs are ch, ck, (qu), ng, sh, th, wh.
6. The letter "y" has been included as a consonant, as a vowel at the end of a two-syllable word where its value is "ē" (some authorities claim its value is "ī") and at the end of a one-syllable word, where its value is "ī".

7. A small number of consonant blends have been included.
8. In summary then, all words included are words which have been encountered frequently by all of the pupils in the control group; all phonetic elements are elements which have been encountered frequently by all pupils by February-March in their reading vocabulary. The validity of these items depends ultimately, on the familiarity of the pupils with them.
9. A summary of the word elements included in the test appears on the next page, together with the frequency with which each element appears in the whole list of words.

SUMMARY OF PHONETIC ELEMENTS, WITH FREQUENCIES

Vowels	Frequency	Consonants	Frequency
⌣ a (closed)	2	b	4
⌣ a (open)	1	c = k	1
a-e	1	c = s	0
aw	1	ck = k	1
all	1	ch	1
aw	1	d	2
ar	1	f	1
ai	1	g (as in go)	1
ay	2	ge, gl = j	0
⌣ e (closed)	4	gh	0
⌣ e (open)	2	gu	0
e = ti	1	h	2
ee	1	j (J)	1
ea	1	k	2
ew	1	l (ll)	2
er	1	m	3
y = ⌣ y = 3	1	n	6
⌣ i (closed)	3	ng	1
⌣ i (open)	1	p	1
i-e	1	qu	1
y = ⌣ y = i	1	r	1
⌣ o (closed)	2	s	2
o = ⌣ o = oo	1	s = z	2

Vowels	Frequency	Consonants	Frequency
oa	1	se = s	1
ow	1	sh	1
ou	1	t	7
oy	1	th	1
or	1	h	2
u (closed)	2	v	1
		w	5
		wh	1
		x	0
		y	1
		z	0
		Consonant Blends	
		— ld	1
		— st	1
		— nt	1
		— lp	1
		pl —	2
		Endings	2
		— s	2
		— ing	1

APPENDIX B
FREQUENCY TABLES



TABLE I
DISTRIBUTION OF SCORES, BY GROUPS AND FOR
COMBINED GROUPS, IN THE WATSON READING
READINESS TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
97.5-.....	5	16	21	158
94.5-97.5.....	5	11	16	137
91.5-94.5.....	5	19	24	121
88.5-91.5.....	4	8	12	97
85.5-88.5.....	3	18	21	85
82.5-85.5.....	3	3	6	64
79.5-82.5.....	4	8	12	58
76.5-79.5.....	3	2	5	46
73.5-76.5.....	4	6	10	41
70.5-73.5.....		4	4	31
67.5-70.5.....		7	7	27
64.5-67.5.....		3	3	20
61.5-64.5.....	1	5	6	17
58.5-61.5.....				
55.5-58.5.....		1	1	11
52.5-55.5.....	3	2	5	10
49.5-52.5.....		2	2	5
46.5-49.5.....				
43.5-46.5.....		1	1	3
40.5-43.5.....	1		1	2
37.5-40.5.....				
34.5-37.5.....		1	1	1

Combined N = 158
N
— = 79
2

Combined Median = 85.5 + $\frac{3(79 - 64)}{2}$
= 85.5 + 2.14
= 87.64

A score of 88 or over is above the median.

TABLE II
DISTRIBUTION OF SCORES, BY GROUPS AND FOR THE
COMBINED GROUP, IN THE *DOMINION TEST OF
LEARNING CAPACITY*

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
51-53.5	2	2	2	158
48-50.5	2	3	5	156
45-47.5	2	7	9	151
42-44.5	5	13	18	142
39-41.5	9	11	20	124
36-38.5	4	14	18	104
33-35.5	4	14	18	86
30-32.5	4	11	15	68
27-29.5	5	13	18	53
24-26.5	2	12	14	35
21-23.5	1	5	6	21
18-20.5	2	3	5	15
15-17.5		3	3	10
12-14.4		6	6	7
9-11.5				
6- 8.5	1		1	1

Combined N = 158
N
— = 79
2

Combined Median = 32.5 + $\frac{3(79 - 68)}{18}$
= 32.5 + 1.83
= 34.33

A score of 35 or over is above the median.

TABLE III
DISTRIBUTION OF TOTAL RAW SCORES, AND FOR THE
COMBINED GROUP, IN THE GATES PRIMARY READING
TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
115-119.5	5	13	18	158
110-114.5	7	9	16	140
105-109.5	7	4	11	124
100-104.5	3	8	11	113
95- 99.5	2	8	10	102
90- 94.5	6	3	9	92
85- 89.5	3	8	11	83
80- 84.5	2	11	13	72
75- 79.5		4	4	59
70- 74.5	1	9	10	55
65- 69.5		6	6	45
60- 64.5	1	9	10	39
55- 59.5		4	4	29
50- 54.5		3	3	25
45- 49.5	1	3	4	22
40- 44.5	2	2	4	18
35- 39.5	1	7	8	14
30- 34.5		3	3	6
25- 29.5		1	1	3
20- 24.5		2	2	2

Combined N = 158
 $\frac{N}{2} = 79$

Combined Median = $84.5 + \frac{5(79 - 72)}{11}$
 $= 84.5 + 3.18$
 $= 87.68$

A score of 88 or over is above the median.

TABLE IV
DISTRIBUTION OF RAW SCORES, BY GROUPS AND BY
COMBINED GROUPS, IN WORD RECOGNITION,
GATES PRIMARY READING TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
48-50.5	5	12	17	158
45-47.5	8	18	26	141
42-44.5	11	9	20	115
39-41.5	8	15	23	95
36-38.5	1	6	7	72
33-35.5	3	12	15	65
30-32.5		3	3	50
27-29.5	1	11	12	47
24-26.5	2	10	12	35
21-23.5	1	3	4	23
18-20.5	1	3	4	19
15-17.5		2	2	15
12-14.5		6	6	13
9-11.5		4	4	7
6- 8.5		3	3	3

Combined N = 158
$$\frac{N}{2} = 79$$

Combined Median = $38.5 + \frac{3(79 - 72)}{23}$
$$= 38.5 + 0.91$$

$$= 39.41$$

A score of 40 or over is above the median.

TABLE V
DISTRIBUTION OF RAW SCORES, BY GROUPS AND BY
COMBINED GROUP, IN SENTENCE READING,
GATES PRIMARY READING TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
45-47.5	4	5	9	158
42-44.5	9	15	24	149
39-41.5	4	6	10	125
36-38.5	4	8	12	115
33-35.5	7	6	13	103
30-32.5	3	13	16	90
27-29.5	1	12	13	74
24-26.5	4	10	14	61
21-23.5		11	11	47
18-20.5		7	7	36
15-17.5		9	9	29
12-14.5	3	6	9	20
9-11.5		6	6	11
6- 8.5	2	2	4	5
3- 5.5		1	1	1

$$\begin{aligned} \text{Combined } N &= 158 \\ \frac{N}{2} &= 79 \end{aligned}$$

$$\begin{aligned} \text{Combined Median} &= 29.5 + \frac{3(79 - 74)}{16} \\ &= 29.5 + 0.94 \\ &= 30.44 \end{aligned}$$

A score of 31 or over is above the median.

TABLE VI
DISTRIBUTION OF RAW SCORES, BY GROUPS AND BY
COMBINED GROUPS, IN PARAGRAPH READING,
GATES PRIMARY READING TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
26	2	6	8	158
25	8	8	16	150
24	4	10	14	134
23	3	8	11	120
22	4	4	8	109
21	3	8	11	101
20	4	6	10	90
19	1	6	7	80
18	2	10	12	73
17	3	6	9	61
16		9	9	52
15	2	6	8	43
14		6	6	35
13	1	4	5	29
12	1	4	5	24
11		8	8	19
10	1	1	2	11
9	1		1	9
8		3	3	8
7		2	2	5
6		2	2	3
5	1		1	1

Combined N = 158
 $\frac{N}{2} = 79$

Combined Median = $18.5 + \frac{1(79 - 73)}{7}$
 $= 18.5 + 0.86$
 $= 19.36$

A score of 20 or over is above the median.

TABLE VII
DISTRIBUTION OF RAW SCORES, BY GROUPS AND BY
COMBINED GROUP, IN THE WORD AND SENTENCE
WRITING TEST

	Experimental Group f	Control Group f	Combined Group f	Cumulative Combined f
40	21	24	45	158
39	5	19	24	113
38	5	14	19	89
37	1	9	10	79
36	1	12	13	60
35	1	2	3	47
34	1	6	7	44
33	2	4	6	37
32	1	4	5	31
31		4	4	26
30	1		1	22
29		2	2	21
28		1	1	19
27		3	3	18
26		1	1	15
25				
24		1	1	14
23	1	5	6	13
22		1	1	7
21				
20		1	1	6
19		1	1	5
18				
17			1	4
16		1	1	3
15		2	2	2

Combined N = 158
N
= 79
2

Combined Median = 37.5 + $\frac{1(79 - 70)}{19}$
= 37.5 + 0.47
= 37.97

A score of 38 or over is above the median.

AN ANALYSIS OF FREE-WRITING IN GRADE IV, V AND VI

FLORENCE I. MCKIE AND M. D. JENKINSON

The Problem

The primary purpose of this study was to analyze the free-writing compositions of grade IV, V, and VI students in city, town, town-rural, and rural schools in north Central Alberta to determine

1. Whether there is a continuous and observable development in students' free-writing through these grades,
2. Whether the free-writing performance of girls is superior to that of boys, and
3. Whether the free-writing of city students is superior to that of town and rural students.

Free-writing, in the context of this investigation, is defined as those compositions written by students who are not directed by specific outlines of content or form but are encouraged to write spontaneously and with minimal restriction about those things, people, or events that have appealed to their interest, imagination, or sensory perception.

The free-writing characteristics selected for analysis included the use of a variety of verbs, sensory impressions, writer reactions, figures of speech and spontaneous expressions, exclamations and questions, and conversation. An analysis of length of composition, development of sentence sense, and use of descriptive or narrative form was also included.

Research that has investigated children's written language is comparatively limited; studies by Edmund (8), Ford (9), and Witty and Martin (18) appear to be forerunners. They have investigated the kind of experiences children write about and the type of characteristic patterns of content and form found in their compositions. The characteristics of free-writing as defined by this study have not been examined by previous research; the research design and analysis of data was therefore explorative.

Research Design

The Sample

Based on a stratified random selection of schools, a total of 1739 free-writing samples was collected. These samples, gathered from sixty-two classrooms in twenty-one schools in Edmonton and surrounding town and rural schools, were taken to be reasonably representative of the intelligence, achievement, and cultural background of Alberta students.

TABLE I
PERCENTAGE DISTRIBUTION OF 1739 FREE-WRITING
SAMPLES

School	Boys		Girls		Total	
	Number	Percent	Number	Percent	Number	Percent
City	359	40.11	343	41.64	702	40.37
Town	219	24.47	193	22.87	412	23.69
Town-rural	245	27.38	253	29.97	498	28.64
Rural	72	8.04	55	6.52	127	7.30
Total	895	51.47	844	48.53	1739	100.00
Grade IV	300	17.25	285	16.39	585	33.64
Grade V	280	16.10	279	16.04	559	32.14
Grade VI	315	18.12	280	16.10	595	34.22
Total	895	51.47	844	48.53	1739	100.00

Table I shows that of the total number of free-writing samples obtained, 40.37 per cent were from city schools; 23.69 per cent from town schools; 28.64 per cent from town-rural schools; and 7.30 per cent from rural schools. The grade distribution was relatively equal with approximately one-third of the samples coming from each grade. The boy-girl distribution in each grade and in city, town, town-rural, and rural schools was also relatively equal.

Collection of Free-Writing Samples

In order to obtain the free-writing samples, the classroom teachers (in the case of city schools, the writer) read an instructional paragraph which was designed to motivate the students to write with minimal limitations about the things they had noticed in their neighborhood. All samples were written between October and December, 1959.

Treatment of Data

Each composition was read by the writer and the various free-writing characteristics were identified and scored separately. The student scores for each characteristic were added to give the total score for each grade, and for the boys and girls of each grade separately. An evaluation grid was designed to credit points for the various free-writing characteristics used by the students and, in this way, to provide a composite free-writing score for the purposes of comparing city, town, and rural school performance. The treatment of data included statistical analysis by means of the "t" of significance at the one per cent and five per cent levels of confidence, and by percentage calculations.

Criteria for Evaluation

Because of the subjective nature of some of the free-writing characteristics analyzed, it was necessary to define these characteristics in rather specific terms. The definitions are illustrated with original examples from the students' free-writing.

Length of Composition

Students who had completed their writing before the allotted time had counted the number of words in their compositions. The remainder were totalled by the writer.

Variety of Verbs

The Dolch Sight Word List was used as the basis for selecting the most common verbs used at the elementary grade level. Only verbs beyond this basic list were scored. For example, "cracked", "bloom", "to experiment", and "peering".

Sensory Impressions

The inclusion of sensory impressions can contribute to the effectiveness of free-writing. Auditory, olfactory, taste, and tactile impressions, and those visual impressions which indicated particularly acute awareness of circumstances or detail were identified. These sensory impressions were credited whether they were direct, using expressions such as "I saw . . .", "I felt . . ." or "I heard . . .", or whether they were implied. For example

There is a pond with the sunlight as ducks and geese go swimming and ducking for food. Beyond there is a church with chapel bells shining in the sun.

As night draws near loons skim over the water with their woeful cries. Owls sing out their mournfull hoo-oo-it-ows as they glide along.

I don't like the city because the buses have a gas that it would make you fell over.

The berrys are called sasgetos. There's purple you know and the red one are soer.

Donld slugged me in the place where I had had my apendix out a month ago.

Writer Reactions

Three fundamental reactions were considered. These were the expression of an opinion, of an emotoin, or of a conscious interest or awareness. Each involves, directly or indirectly, a relationship between the student writer and what he is writing about. For example

I think Id rather live on the farm than live in the city because on the farm you can have more pets.

I watch a program on T.V. Which was a hair-raising story in some parts . . . So I thought of something funny so I wasn't scared.

She is a tomboy, and follows my borthers' fried Wyane all over the place. She plays football with us every time we do and walways makes a touch-down. And she also plays baseball and makes a homerun. Andof course plays pome-pome-pullaway and throw me it I try to get her And yet she is very pretty in a dress.

Effective Forms of Expression

Younger children have a spontaneity in their speech that gives charm to their oral expression. To measure this type of uninhibited expression, four areas of effective expression were considered.

1. Figures of speech identified with similes, metaphors, and personification.
2. Spontaneous expressions were identified as casual, informal, or uninhibited comments of enthusiasm, doubt, or reflection which suggested a conversational interchange between the writer and the reader. For example

Let me take you back four years.

Boy, did Mom give me heck!

I guess you would consider our neighborhood in the part of the spoiler part of town. Us kids don't mind.

3. Conversation appeared in two characteristic forms, indirect and direct, each revealing an increasing maturity in the reporting of the spoken word and an increasing ability and confidence in the use of quotation marks. For example

The man said that I'd have to pay him five dollars if I wanted the dog.

Then I said you are crool mrs. Jones. If you do that once more i will tell the poleac.

"I rieplied to give rides!"

"She rieplied oh yes!"

"Look at this one," said Joanne. "Its for me.

I wonder whats' in it," exclaimed Lloyd.

4. Sentence Variety. Exclamations and questions were taken as the most obvious examples of sentence variation. These were credited as "implied" with no specialized punctuation, or as "correct" with proper punctuation. For example

That was a close call.

Kerplosh! He had fallen in!

But who was coming.

What do you think I saw?

Sentence Sense

While on the whole this study emphasized the positive aspects of children's writing, a brief analysis of sentence sense showed that maturity in the use of sentences does vary. Characteristic sentence immaturities examined were fragmentary sentences, run-on sentences, unvaried sentence beginnings, unvaried use of the first person "I", and itemized writing. The basic criteria for identifying these sentence characteristics was that three or more sentences of any one type must be found in any one composition. A single composition may have several of these characteristic sentence faults and was therefore scored under each category.

Composition Form

The form of composition was considered as "narrative", that is, relating an experience in sequential form and thereby having a story structure, or as "descriptive", describing or discussing some person, object, or animal in a manner which explains or illuminates but does not involve experiential events.

TABLE II
SUMMARY OF THE MEANS FOR FREE-WRITING
CHARACTERISTICS FOR THE TOTAL SAMPLE

Free-Writing Characteristic	Mean	Standard Deviation	Standard Error	t-score	Level of Significance
Length of Composition					
Grade IV.....	88.35	31.05	} 2.29	4.72	1%*
V.....	99.15	44.99			
VI.....	137.55	47.91			
Variety of Verbs					
Grade IV.....	2.46	2.13	} .13	3.69	1%
V.....	2.94	2.31			
VI.....	4.12	3.21			
Sensory Impressions					
Grade IV.....	.43	.62	} .02	5.83	1%
V.....	.57	.85			
VI.....	.87	1.01			
Writer Reactions					
Grade IV.....	1.18	1.34	} .08	4.51	1%
V.....	1.55	1.45			
VI.....	2.28	1.94			
Composite Scores					
Grade IV.....	5.04	3.85	} .03	23.33	1%
V.....	5.74	5.18			
VI.....	8.33	7.14			

*A significant increase between grade IV and grade V students, in favor of grade V.
†A significant increase between grade V and grade VI students, in favor of grade VI.

TABLE III
SUMMARY OF 1,739 STUDENTS' FREE-WRITING
PERFORMANCE IN GRADES IV, V, AND VI

Free-Writing Characteristic	Trend from Grade IV to VI	Percentage of Students Contributing	Superior Performance*			
			IV	V	VI	Total
Lenth of Composition	Significant increase	100.00	NSD	G5%	G5%	—
Variety of Verbs	Significant increase	85.62	G5%	B1%	NSD	—
Sensory Impressions	Significant increase	34.27	B1%	G1%	NSD	—
Writer Reactions	Significant increase	70.44	G1%	G1%	G1%	—
Figures of speech	Slight increase	1.73	G.	G.	G.	G.
Spontaneous Expressions	Decrease	6.50	G.	B.	G.	G.
Exclamations	Increase	5.12	G.	G.	G.	G.
Questions	Approx. equal	2.13	G.	G.	B.	=
Indirect Conversation	Approx. equal	15.47	G.	G.	B.	G.
Direct Conversation	Approx. equal	6.67	G.	G.	G.	G.
Sentence Sense	Decrease of Sentence faults	26.91	G.	G.	G.	G.
Composite Free- Writing Scores	Signifite increase	100.00	G1%	G1%	NSD	—
*Boys is B; Girls is G; Equal Performance is =; No Significant difference is NSD.						

Analysis of Data

Table II shows that the increase from grade IV to grade VI in length of composition, variety of verbs, sensory impressions, writer reactions, and the composite free-writing scores was significant at the one per cent level of confidence. This indicated a steady and increasing development in these free-writing characteristics as the students matured. On this basis, there is no language plateau in grades IV, V, and VI in free-writing. However, it is important to note that the mean number of verbs, sensory impressions, and writer reactions is extremely small, and represents, in some cases, only a half word gain from one grade to the next. While statistically sig-

nificant, such increase cannot be accepted as satisfactory from the educational point of view.

Table III shows that the pattern of development for the remaining free-writing characteristics from grade IV to grade VI varied. For example, while the use of exclamations showed an increase throughout the grades, the use of spontaneous expressions decreased. The use of questions and conversation remained approximately equal. It is significant, however, that there was a decrease in the number of students making sentence faults. Table III further summarizes the difference in free-writing performance between the boys and

TABLE IV
SUMMARY OF FREE-WRITING CHARACTERISTICS
CONTRIBUTED BY THE TOTAL SAMPLE

Free-Writing Characteristic	Number of Examples	Percentage Classifica- tion	Percentage Boys	Distribution Girls
Variety of Verbs	5562	100.00	52.04	47.96
Sensory Impressions	1067	100.00	49.29	50.71
Auditory	279	25.95	11.71	14.24
Olfactory	22	2.05	1.02	1.02
Taste	10	.93	.28	.65
Tactile	353	32.83	19.53	13.30
Visual	403	37.48	16.65	20.83
Writer Reactions	1283	100.00	43.62	56.38
Figures of Speech	57	100.00	47.37	52.63
Simile	45	78.94	36.79	41.15
Metaphor	10	17.54	5.26	12.28
Personification	3	5.22	3.47	1.75
Spontaneous Expressions	134	100.00	43.28	56.72
Exclamations	109	100.00	40.35	59.65
Implied	70	64.20	25.68	38.51
Punctuated	39	35.80	14.67	21.14
Questions	41	100.00	51.22	48.78
Implied	21	46.35	29.27	17.07
Punctuated	20	53.65	21.95	31.71
Indirect Conversation	463	100.00	45.57	54.43
Direct Conversation	178	100.00	39.44	60.66
Implied	116	65.19	26.97	38.22
Attempted	25	14.01	2.25	11.76
Punctuated	37	20.79	10.12	10.68
Sentence Faults	516	100.00	62.98	37.02
Fragmentary	20	3.86	2.89	.97
Run-on	333	64.27	39.95	24.32
Unvaried Beginnings	163	31.46	19.88	11.58
Repetitive "I"	17	100.00	47.06	52.94
Itemized Writing	35	100.00	41.13	58.87
Descriptive Form	1156	66.47	32.60	33.87
Narrative Form	583	33.53	19.31	14.22

TABLE V
SUMMARY OF SCHOOL PERFORMANCE AS MEASURED BY
THE MEAN COMPOSITE FREE-WRITING SCORES FOR THE
TOTAL SAMPLE

Schools	Number of Students	Range of Points	Mean Points	Standard Deviation
City	702	0-29	5.52	4.05
Town	412	0-37	7.51	4.02
Town-rural	498	0-33	7.80	4.68
Rural	127	0-25	6.01	5.36
Total	1739	0-37	6.66	7.30

TABLE VI
COMPARISON OF SCHOOL PERFORMANCE MEASURED BY
THE MEAN COMPOSITE FREE-WRITING SCORES FOR THE
TOTAL SAMPLE*

Schools Compared	SE _d	t-score	Level of Confidence
City and Town*	.24	8.29	1%
City and Town-rural*	.25	9.12	1%
City and Rural	.50	.98	NSD
City and Total Sample*	.31	3.68	1%
Town and Town-rural	.28	1.12	NSD
Town* and Rural	.51	2.94	1%
Town* and Total Sample	.33	2.58	1%
Town-rural* and Rural	.51	3.51	1%
Town-rural* and Total Sample	.34	3.35	1%
Rural and Total Sample	.55	1.18	NSD

*School with the greater mean composite free-writing score.

girls at each grade level. In the majority of cases, the girls gave a superior performance to that of the boys. There was, however, no significant difference between the boys and girls in the total composite free-writing score at the grade VI level.

Table IV summarizes the total number of examples identified and categorized for each free-writing characteristic and indicates, where applicable, the classification of the characteristic in terms of implied or correct punctuation or types of examples. For example, of the total 178 examples of direct conversation identified, two thirds were

written without any specialized punctuation, while twenty per cent were correctly punctuated. Table IV also indicates the percentage of examples contributed by boys and by girls. For example, two-thirds of the sentence faults were made by boys; almost two-thirds of the examples of direct conversation were made by the girls.

A comparison of the mean composite free-writing scores for the different schools participating in the study is shown in Tables V and VI. The analysis of data showed that city schools scored significantly lower than town and town-rural schools and significantly lower than the total sample.

Conclusions

1. There was a continuous and observable development in the free-writing characteristics analyzed from grade IV to grade VI. The increase in length of composition, in use of variety of verbs, sensory impressions, writer reactions, and in composite free-writing scores was significant at the one per cent level of confidence from grade IV to grade VI. Other free-writing characteristics measured by percentage calculations only, such as figures of speech, conversation, and sentence sense, showed generally positive development throughout the grades. Specifically, the pattern of development for the various free-writing characteristics was as follows:
 - (a) The mean number of words for each grade level was 88.35, 99.15, and 137.55 words. There was a greater increase in length of compositions from grade V to grade VI, than from grade IV to grade V. This indicates a steady and increasing growth in the ability to write longer compositions as the students mature.
 - (b) The number of verbs beyond the Dolch List showed a significant increase from grade IV to grade VI. The mean number of verbs per composition was extremely small, ranging from 2.46 in grade IV to 4.12 in grade VI. The difference between boys and girls showed no consistent pattern.
 - (c) There was no linear development in the use of figures of speech in grade IV, V, and VI. Similies were most commonly used and showed an increasing development from grade IV to grade VI. Metaphors and personification were few and most frequently used by grade V students. More girls used figures of speech than did boys.
 - (d) Spontaneous expressions were used twice as frequently by grade IV students, as by grade V and VI students.
 - (e) Although the mean number of sensory impressions was less than one per composition in all grades, one-third of all the

students in the sample used sensory impressions in the free-writing. Sensory impressions therefore represent a slight, but potential characteristic in the free-writing compositions analyzed. Of the total number of sensory impressions categorized, tactile and visual impressions each accounted for a third. Auditory impressions accounted for somewhat less than one third, while olfactory and taste impressions were infrequent. Throughout the grades a greater and increasing percentage of boys used tactile impressions, while a greater and increasing percentage of girls used visual impressions.

- (f) The number of writer reactions increased significantly from grade IV to grade VI. However, very few writer-reactions were included, the mean number per students being 1.18 in grade IV and 2.28 in grade VI. The difference between boys and girls was significant in favor of the girls at all grade levels. There appeared to be a development in the type of writer reactions given as the students matured. The simple reactions of personal approval or disapproval were expanded to more reflective and insightful comments.
- (g) Students at this age level showed only a slight tendency to use direct and indirect conversation in their free-writing. Approximately twice as many students used the indirect form as used the direct form. The differences between the grades were negligible. The use of conversation in free-writing at this age level appeared to be largely an incidental rather than learned technique. Knowledge and correct use of punctuation for direct conversation were relatively limited. Two-thirds of the direct conversation samples were written without specialized punctuation.
- (h) The analysis of writer reactions and sentence variety, particularly the use of the repetitive "I", indicated that the students of grades IV used more ego-centric comments and more sentences beginning with "I" than did students of grade V and grade VI.
- (i) Approximately five per cent of the total sample used exclamations as one form of sentence variation in their free-writing. In all grades, more girls used exclamations than did boys. Questions were used infrequently also.
- (j) Approximately one quarter of the total number of students in the sample made sentence faults of one kind or another. The percentage of students making sentence faults decreased from grade IV to grade VI, indicating that run-on sentences, fragmentary sentences, and unvaried sentence beginnings

were being gradually eliminated. In all grades, more boys made sentence faults than did girls. The run-on sentence was the most common fault, accounting for two-thirds of the total number of sentence faults identified. Sentences with unvaried beginnings accounted for slightly less than one-third, while fragmentary sentences made up approximately four per cent of the total number of sentence faults.

- (k) The students in grade IV, V, and VI showed an increasing preference for descriptive rather than narrative compositions. Two-thirds of the total free-writing compositions were descriptive. A greater percentage of girls wrote descriptive compositions, while boys tended to write more narrative compositions.
- 2. The free-writing performance of girls was superior to that of the boys. The difference was significant at the grade IV and grade V level. However, the difference between boys and girls at the grade VI level was not significant. The language maturation of boys seemed somewhat later than that of the girls, but as the students matured, the difference between boys and girls in written language decreased.
- 3. The free-writing performance of city students was not superior to that of town and rural students. On the basis of the composite free-writing scores, town students were superior to both city and rural students. There was no significant difference between city and rural students in their free-writing performance.

Limitations

The conclusions of this study are governed by certain limitations inherent in the research design and analysis of data.

- 1. The total sample of students was assumed to be representative of the grade IV, V, and VI students in north central Alberta. However, the proportion of city, town, town-rural, and rural students may not be precisely that of the population as a whole.
- 2. The directed-motivation presented to the students for the purpose of obtaining the free-writing samples may have been more appropriate for town and rural students than for city students.
- 3. It may have been a disadvantage for city students to have been required to write for a totally unfamiliar person (the writer) rather than for their classroom teacher.
- 4. The various free-writing characteristics selected for investigation, the criteria of evaluation, and the treatment of data were arbitrarily defined by the writer. The conclusions of this study are limited by the degree to which these free-writing character-

istics are accepted as legitimate components of effective expression, the appropriateness of the criteria of evaluation, and the accuracy of the analysis of data.

5. The subjective nature of many of the free-writing characteristics lends the possibility of variation in judgment regarding the identification and classification of the samples.
6. Further, the use of an evaluation grid to present a composite free-writing score may or may not be regarded as a valid approach to the problem of assessing the quality of free-writing.

BIBLIOGRAPHY

1. Burrows, A. T., D. C. Ferebee, et al. *They All Want To Write*. New York: Prentice-Hall Inc., 1952.
2. Carlson, Ruth Kearney. "Seventeen Qualities of Original Writing." *Elementary English*, XXXVIII, December, 1961, pp. 576-79.
3. Conquest, George Rienzi. "A Survey of English Language Achievement in Grades Four and Seven in Selected Alberta Schools." Unpublished Master's thesis, The University of Alberta, Edmonton, Alberta, 1954.
4. Coutts, H. T., and H. S. Baker. "A Study of the Written Composition of a Representative Sample of Alberta Grade Four and Grade Seven Pupils." *The Alberta Journal of Educational Research*, I, June, 1955, pp. 5-18.
5. Department of Education. *Elementary Language Bulletin 2C*. Edmonton: The Queen's Printer, 1959.
6. Dunlop, G. M., S. Hunka, and H. Zingle. "Individual Differences in Alberta Schools." *The Alberta Journal of Educational Research*. I. December, 1955, pp. 5-14.
7. Dusel, William J. "Some Semantic Implications of Theme Corrections." *The English Journal*, XLIV, October, 1955, pp. 390-97.
8. Edmund, Neal R. "A Study of the Relationship Between Prior Experiences and the Quality of Creative Writing Done by Seventh-Grade Pupils." *Journal of Educational Research*, LI, March, 1959, pp. 481-92.
9. Ford, C. T. "Developments in Written Composition During the Primary School Period." *British Journal of Educational Psychology*, XXIV, February, 1954, pp. 38-45.
10. Gray, Mary Edith. "An Analysis of Language Themes of Grade Five, Grade Eight, and Grade Eleven." Unpublished Master's thesis, The University of Alberta, Edmonton, Alberta, 1959.
11. Hourd, Marjorie L. *The Education of the Poetic Spirit*. London: William Heinemann Ltd., 1949.
12. Loban, Walter, *Language Ability in the Middle Grades of the Elementary School*. Department of Education, University of California, 1961.
13. Russell, David H. *Children's Thinking*. New York: Ginn and Co., 1956.
14. Smith, Dora V. "Growth in Language Power as Related to Child Development." *Teaching Language in the Elementary School*. The Forty-Third Yearbook of the National Society for the Study of Education, Part II. Nelson B. Henry (ed). Chicago: The University of Chicago Press, 1944, pp. 52-97.
15. Strickland, Ruth G. *The Language of Elementary School Children*. Its Relationship to the Language of Reading Textbooks and the Quality of

- Reading of Selected Children. Volume XXXVIII of Bulletin of the School of Education, Indiana University, 1962.
16. Watts, A. F. *The Language and Mental Development of Children: An Essay in Educational Psychology*. Toronto: George G. Harrap and Co. Ltd., 19'0.
 17. Wiseman, Stephen. "The Marking of English Compositions in Grammar School Selections." *The British Journal of Educational Psychology*, XIX, Part III, November, 1949, pp. 2000-09.
 18. Witty, Paul and William Martin. "An Analysis of Children's Compositions Written in Response to a Film." *Elementary English*, XXXVI, March, 1957, pp. 158-63.

SOME FACTORS AND ACHIEVEMENT IN LITERATURE 20

JOHN HAWRELKO

The Problem and Method

The five inspectorates—Bonnyville, Lac La Biche, Smoky Lake, St. Paul, and Thorhild—cover a fairly large area of northeastern Alberta. Nearly twenty distinct racial groups engaged in a predominantly agricultural pursuits are settled here. Understandably, then, many children in this area came from a variety of economic and social levels in both urban and rural environments. Undoubtedly, these factors exert no small influence on student achievement generally; and in subject areas as in the language arts these factors may be of considerable importance.

This study, therefore, was undertaken to investigate the relationship between achievement in Literature 20 of students of three predominant racial stocks (English, French, and Ukrainian) and each of the following factors:

- (a) the number of books read,
- (b) the quantity of periodicals the student had access to,
and/or
- (c) the size of the home library.

In addition, differences in achievement in Literature 20 between students of English racial stock and French racial stock, and between students of French racial stock and Ukrainian racial stock were investigated statistically. This study, also, investigated the achievement of rural or "country" students in Literature 20 and the achievement of urban or "town" students.

The method of analysis of the data obtained from a questionnaire and an examination in Literature 20 involved the use of non-parametric statistics: the median, and the extension of the median test; the Kruskal-Wallis test; and the Mann-Whitney U test.

Review of Related Theory and Research

Statements on literature teaching in general

Aims in education do change, often from generation to generation, reflecting goals of personal and social development based on the needs, problems and interests of the student. In Alberta the primary aim of the secondary school today is to assist adolescent growth toward maximum self-realization, involving both emotional and intellectual aspects of the individual's life. Among the adolescent's

intellectual potentials the following are said to be desirable ends (12):

- ability to think rationally, to express thoughts clearly and to read and listen with understanding
- an understanding and appreciation of cultural heritage
- the development of suitable recreational and leisure time activities

Through literature, specifically, Loban, Ryan, and Squire (16) define these as goals:

1. Literature for self understanding and for arriving at values of life.
2. Literature as imaginative illumination and the recognition of beauty.
3. Literature for a balanced perspective in life.

The objectives listed above are basic to the development of an adequate literature program for the secondary school in Alberta. Knowledgeable people in the language arts believe that through literature many personal and social needs can be satisfied. Heaton and Lewis (14), for example, believe that:

Biography, fiction, and drama offer readers an opportunity to identify emotionally with human beings who are in interaction with their fellows. They provide access to the feelings of other people in a way otherwise offered only by face-to-face contacts.

The high school is no longer an institution intended primarily to provide training for college-bound students. It has assumed an additional role—that of providing students, some of whom have abilities which are more limited, with a training for life. Training for life thus presupposes training through the satisfaction of individual needs. Since our schools are essentially reading schools where much is dependent on one's reading ability, and since there is a wide range in reading ability at any one level we can hardly over-emphasize the importance of adequate reading and literature programs tailored to individual needs.

Yet reading skill, indispensable as it may be, is not all that is required to ensure success in a literature program. McCullough (19) speaks of an "appreciation factor" apparently independent of reading skill. Loban, Ryan, and Squire enlarge on this further:

The more subtle challenges of literature lie beyond fundamental skills of comprehending simple prose structures. Like all fine arts, literature uses special ways to evoke experience in others. By particular uses and arrangements of words, it seeks to express realms of experience inaccessible to ordinary language Without training and imagination they [students] cannot understand the specific language of literature nor enter the domain it charts.

Thus appreciation that exists apart from comprehension in literature is probably a product of special skills and attitudes which are required of the reader of literature material.

Burton (2) has experimentally verified the existence of an appreciation factor. He concludes his study of 190 Minnesota high school students by saying that "high intelligence and silent reading

ability applied to literary materials will not guarantee literary appreciation", though these often emerge as important factors.

Statements on the influence of books, home libraries, and periodicals on students

Researchers and educators have long recognized the importance of books and periodicals in the lives of high school students. Cass (5) and Chase (6) believe that books remain "the richest source of help other than life." The American Association of School Librarians (1) believes that creative, informed and wise people will be those who will use library resources frequently. Contrary to popular belief and in spite of the incursion of TV and the automobile young people, Fay insists (11), can and do read, though an Ontario study (8) showed that home libraries were not at all adequate. The Ontario committee believes that recreational reading is "above all the responsibility of the parents." Sheldon and Carillo (24), studying the influence of the parents and the home on children's reading concluded that:

As the number of books in the home increases, the per cent of good readers increases and the per cent of average and poor readers decreases.

In the matter of periodical reading much has been written about the influence of this on adolescents both in and out of school. The National Society for the Study of Education devoted its entire 1954 yearbook to mass communication theory and research (4). It makes a plea to the school to lay a basis for proper and intelligent use of the mass media.

Statements on bilingualism

Man has always been curious about the phenomenon of language but not until recent times has language, *per se*, undergone the scrutiny of linguistic scientists. Soffietti (27), an expert in this area, for example, insists that mere acquisition of "linguistic habits" hardly defines the polyglot, and that one must also regard the bilingual as a product of a particular cultural milieu. He proceeds thus to classify the bilingual into four distinct types recognizing the bilingual's cultural influence. Hitherto studies in bilingualism have been of concern to psychologists, sociologists, and educators only in the aspects which determine bilingualism, or through which its effects were being defined. Virtually dozens of studies of this type have been completed in recent decades. Of considerable interest, in this regard, is the study carried out by Darcy (10) in which she investigates the effect of bilingualism upon measurement of the intelligence of Italian children of pre-school age in New York. In the United Kingdom Jones (15) had conducted somewhat similar experiments in which he was primarily concerned with the language handicaps of Welsh-speaking children.

In Alberta Robinson (23), Sullivan (28), and Skuba (26) identified difficulties that derived from the bilingual background of the students. Reid (22), Coull (9), and Olson (20) conducted research studies in which bilingualism was important.

In recent times, however, linguistic experts such as Mackey (17) write that mere identification of bilingualism is not enough if we are to know the influence of bilingualism in the life of the polyglot. He suggests that, preliminarily, studies should be conducted in an effort to describe bilingualism: to find out how bilingual a person is—in speech and writing, and in comprehension and expression of each language. In this description attention should be directed to the function of the languages. Studies should determine how, why, and to what extent the bilingual alternates between his languages. The study, also, should include the question of the amount of interlingual interference. These four aspects suggest a framework of research, according to Mackey, which is necessary if we are to gain accurate conclusions in reference to the effects of bilingualism on the individual.

Statements on the differences between rural and urban student achievement

There is a considerable difference of opinion regarding the definition of a rural area as contrasted to an urban community. There is a sharp divergence of meaning between what is defined as "rural" in the *Encyclopedia of Educational Research*, and what is regularly and loosely defined as rural in Alberta. Therefore studies done outside this province are not strictly relevant to this investigation. However, Martens (18) carried out a research resembling this one in which he studied the educational achievement of eighth-grade pupils in one room rural and graded town schools in Iowa. His research led him to conclude that "town" pupils outperformed their rural counterparts in all areas of learning but one. The one exception being spelling. In Alberta, among recent researchers, Pritchard (21), Climenhaga (7), and Carmichael (3) have conducted studies involving in varying degrees this rural-urban comparison of pupil achievement. Urban pupils seemed to have been better achievers.

Plan, Procedure, and Collection of Data

The subjects of this study consisted of 607 Literature 20 students in the inspectorates mentioned earlier. These students were given an examination designed to test for the purposes and aims that were set out in the *Curriculum Guide*. The investigator however, recognized the very limiting aspects of such a testing device in view of the fact that in a literature test it is often difficult to identify, de-

fine, and control the variables involved. With the examination framed for this study, and attached to it, a student and teacher questionnaire was also sent out through which certain other required information could be obtained.

Analysis of the data

Nonparametric statistics were used to analyze the data. Some of the data were cast in contingency tables and the probabilities of the distribution were calculated by the chi-square. The interpretations were based on the value of p. The value of p greater than .05 in this study was interpreted as non significant, indicating that there appeared to be no significant difference between the groups compared for the particular item under investigation. A value of p less than .01 was interpreted as very significant.

I. Analysis of factors in the problem

A. The relationship of books read by the students of the three racial stocks and achievement in Literature 20.

The statistical test, the median test, was selected to test the null hypothesis for each of the racial stock groups relative to the number of books read. The significance level was set at: $\chi = .05$.

TABLE I
STUDENTS WHO READ TEN BOOKS OR LESS COMPARED
WITH THOSE WHO READ ELEVEN BOOKS OR MORE IN
THE ENGLISH RACIAL STOCK GROUP

Raw Scores Relative to Median	No. of Books Read		Totals
	Group I (10 books or less)	Group II (11 books or more)	
No. of scores that exceed the combined median	17	31	48
No. of scores that are at the combined median or below	19	26	45
Totals	36	57	93

Combined Md. = 72.5; $\text{Chi}^2 = .31$; $\text{df} = 1$; $.35 < p > .25$

Since p yielded by the test was greater than $\chi = .05$ the null hypothesis—that there was no significant difference between the median scores of students who read ten books or less, and those who read eleven books or more—was accepted.

For the remaining racial stocks the tables are left out for purposes of conserving space. Only the data are given below:

2. For the French racial stock group the following is a summary of the data:

Combined Md. = 64.5; $\chi^2 = 1.22$; $df = 1$; $.15 < p > .10$

Since p yielded by the test was greater than $\chi = .05$ the null hypothesis was accepted.

3. For the Ukrainian racial stock group the following is a summary of the data:

Combined = 64.5; $\chi^2 = 8.95$; $df = 1$; $.005 < p > .005$

Since p yielded by the test was less than $\chi = .05$ the null hypothesis was rejected.

B. Again, for purposes of brevity the results of the following factors are combined and charted below:

1. the relationship of the number of periodicals read by the students of the three racial stocks and achievement in Literature 20, and,
2. the relationship of the size of the home library of students in the three racial stocks and achievement in Literature 20:

Group	No. of periodicals read	Size of home library
English	no relationship	relationship
French	no relationship	no relationship
Ukrainian	no relationship	no relationship

II. Analysis of other factors in the problem

A. In respect to the significance of the difference in achievement in literature 20 among the students of the three racial stocks, the data relative to this problem was subjected to the Kruskal-Wallis test for k independent samples (13 and 25).

The following data were used:

Ukrainian	= n_3	= 284	—	R_1	= 27295
English	= n_1	= 93	—	R_2	= 28897.5
French	= n_1	= 122	—	R_3	= 68553.5

The H yielded by the test was 10.31. For df of 2 and the significance level set at $\chi = .05$ the probability was: $.01 < p > .001$. Thus the null hypothesis was rejected. The three samples were not drawn from populations whose scores had the same median.

B. Since the Kruskal-Wallis test emphatically rejected the null hypothesis that the k independent samples were drawn from the

same population, did the students of the English racial stock achieve significantly higher? (The median score in the test for the English racial stock group was 72.0 compared to the French group's score of 63.5, and the Ukrainian group's score of 64.5). The Mann-Whitney U test was used with the following data and with the following results:

1. English-French comparison

$n_1 = 93$ = the number of English students
 $n_2 = 122$ = the number of French students
 $R_1 = 11379$; $R_2 = 11841$; $\Sigma T = 259$; $U = 4338$
 $Z = -2.99$; $p = .0014$; null hypothesis rejected

2. English-Ukrainian comparison

$n_1 = 93$ = the number of English students
 $n_2 = 122$ = the number of French students
 $R_1 = 20288$; $R_2 = 50970$; $\Sigma T = 978$; $U = 10495$
 $Z = -3.78$; $p < .00011$; null hypothesis rejected

3. French-Ukrainian comparison

$n_1 = 122$ = the number of French students
 $n_2 = 284$ = the number of Ukrainian students
 $R_1 = 18221$; $R_2 = 58060$; $\Sigma T = 1451$; $U = 13930$
 $Z = 1.4842$; $p = .0694$; null hypothesis accepted

In summary, analysis of the data indicated that students of English racial stock achieved significantly higher in Literature 20 than either the French or Ukrainian students in the same geographical area. There was no significant difference in achievement in Literature 20 between the French and Ukrainian students.

C. In respect to the significance of the difference of achievement in Literature 20 between the rural and urban student in this geographical area the following data were used.

n. of urban students	206
n. of rural students	396
median score for urban students	73.0
median score for rural students	63.0
combined median	65.5
Chi ²	19.9
df	1
p	<.0005

The median test was selected to test the null hypothesis—that there was no significant difference between the medians of the scores of urban and rural students in the testing area.

Conclusions

1. From the analysis of the data it would appear that Ukrainian students who read more books in their leisure hours achieved higher in Literature 20.
2. There appeared to be no immediate relationship between the accessibility of periodicals and achievement in Literature 20.
3. For the English racial group, only, achievement in Literature 20 was related to the size of the home library.
4. Analysis of the data indicated that students of English racial stock achieved significantly higher in Literature 20 than either the French or Ukrainian students in the same geographical area.
5. Urban students in the geographical area under study achieved significantly higher in Literature 20 than did the rural students in the same area.

REFERENCES

1. American Library Association. *Standards for School Libraries. A report by the American Association of School Librarians.* New York: American Library Association, 1960.
2. Burton, Dwight L. "The Relationship of Library Appreciation to Certain Measurable Factors." *Journal of Educational Psychology*, 43:436-443, November, 1952.
3. Carmichael, A. "A Survey of the Achievement of Alberta School Children in Reading." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1954.
4. Dale, E. (ed.) *Mass Media of Communications*, XLV Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1954.
5. Cass, J. (ed.) *Books in Schools.* New York: American Book Publisher's Council, Inc., 1961.
6. See reference 4. above.
7. Climenhaga, C. E. "A Survey of Arithmetical Achievement of Grade Eight Pupils in Alberta Schools." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1955.
8. Committee on Children's Recreational Reading in Ontario. "The Recreational Reading Habits of Ontario School Children." *Study Pamphlets in Canadian Education* No. 10. Toronto: The Copp Clark Publishing Company, Ltd., 1952.
9. Coull, W. M. "A Formative Survey of Reading of Alberta School Children in Relation to Intelligence, Sex, Bilingualism and Grade Placement." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1956.
10. Darcy, N. T. "The Effect of Bilingualism on the Measurement of the Intelligence of Preschool age." *Journal of Educational Psychology*, 37:22-44, January, 1946.
11. Fay, Leo. C. *Reading in High Schools.* Washington: National Education Association, 1956.
12. Government of Alberta. *Senior High School Curriculum Guide for English.* Edmonton: Government of Alberta, Department of Education, 1957.
13. Guilford, J. P. *Fundamental Statistics.* New York: McGraw-Hill Book Company, 1956.

14. Heaton, M. M., H. B. Lewis. *Reading Ladders to Human Relations*. Washington: American Council of Education, 1955.
15. Jones, W. R. "Influence of Reading Ability in English on the Intelligence Test Scores of Welsh-speaking Children." *British Journal of Educational Psychology*, 23:114-120, June, 1953.
16. Loban, W., M. Ryan, J. Squire. *Teaching Language and Literature*. New York: Harcourt Brace, and World Inc., 1961.
17. Mackey, W. G. "The Description of Bilingualism." *The Canadian Journal of Linguistics*, 7:2:51-85, Spring, 1962.
18. Martens, C. C. "Educational Achievement of Eighth-Grade Pupils in One-room and Graded Town Schools." *Elementary School Journal*, 54:523-525, May, 1954.
19. National Conference on Research in English. *What We Know About High School Reading: A Report Prepared by the N.C.R.E.* Chicago: National Conference on Research in English, 1955.
20. Olson, D. R. "The Influence of Foreign Language Background on Performance on Selected Intelligence Tests." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1962.
21. Pritchard, R. O. "A Survey of the Arithmetical Achievement of Grade Five Pupils in Alberta Schools." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1955.
22. Reid, T. J. "A Survey of the Language Achievement of Alberta School Children in Relation to Bilingualism, Sex, and Intelligence." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1954.
23. Robinson, C. H. "A Study of the Written Language Errors of 1238 Pupils of Ukrainian Origin." Unpublished Bachelor's Thesis, The University of Alberta, Edmonton, 1934.
24. Sheldon, W., L. Carillo. "Relation of Parents, Home, and Certain Development Characteristics to Children's Reading Ability." *Elementary School Journal*, 52:262-270, January, 1952.
25. Siegel, Sidney. *Nonparametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill Book Company, 1956.
26. Skuba, M. "An Analysis of English Errors and Difficulties Among Grade Ten Students in the Smoky Lake School Division." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1955.
27. Soffiatti, J. P. "Bilingualism and Biculturalism." *Journal of Educational Psychology*, 46:222-226, April, 1955.
28. Sullivan, D. M. "An Investigation of the English Disabilities of Ukrainian and Polish Students in Grade IX, X, XI, XII of Alberta Schools." Unpublished Master's Thesis, The University of Alberta, Edmonton, 1946.

CASE STUDIES OF SOCIALLY ISOLATED MALES IN SENIOR HIGH SCHOOL

J. K. NIELSEN

When an individual's acquaintances rarely or never desire his company, that individual may be described either as socially isolated or as socially unaccepted. Depending on the circumstances, being socially isolated may have both advantages and disadvantages for the isolated, and for society.. For example, in the opinion of L. H. Meek (2), academic underachievement is a natural consequence of social isolation, and it may, according to Bloch (1), develop into schizophrenia.

When social isolation appears to be harmful to the isolate or to society, the desirability of its reduction or elimination is usually considered. Our efforts to reduce social isolation are likely to become more effective as our knowledge of the factors involved increases. Considerable research dealing with the general problems of social acceptance is already available, but research dealing specifically with socially isolated male senior high school students is difficult to find. The problem of this study was to identify factors exclusive of race, color, and school attendance, which were characteristic of some socially isolated high school boys. No attempt was made to test the validity of any hypothesis.

Selection of Cases

A "three-choice, three-criterion" sociometric test was administered to the eleven classes of grade eleven students in Victoria Composite High School in Edmonton, Alberta. The students were twice asked to name the three persons they most preferred as companions in three different situations. In the first listing they were asked to confine their choices to other students in their classroom. They were then asked to make a second selection, this time the only restriction being that they confine their choices to residents of Edmonton. In obtaining a measure of sociometric status for each student, first choices were given a weight of three, second choices a weight of two, and third choices a weight of one.

For the purposes of this study, a male classroom social isolate was defined as any boy who received less than two fifths of the average choice receipts per male class member. Sixty-six of the 246 boys in grade eleven were by this definition designated as classroom isolates. Since race, color, and school attendance could have played a part in the isolation of twenty-four boys, the group

under consideration was reduced to forty-two classroom social isolates. For each of these forty-two boys a second sociometric score (All Edmonton score) was obtained in the following manner. Those persons who were chosen by the classroom social isolate when his choices had been limited to residents of Edmonton were in turn requested to fill out a "three-choice, three criterion" sociometric test also limiting their choices to people in Edmonton. To the weighted choices given to an isolate by persons he chose who were not in grade eleven in his school, were added the weighted choices given to him by grade eleven students in his school when their choices had been limited to Edmonton residents. The classroom isolates were then ranked in order on the basis of this "All Edmonton" sociometric score, and, starting with the boy receiving the lowest score, each was invited individually by the investigator to take part in the study. This process was continued until fourteen boys had agreed to take part. Since the twelfth boy was absent from school during the selection of cases, he was not invited. However, all boys who were invited agreed to cooperate.

Study of Cases

After the cases had been selected, their cumulative record cards were examined for pertinent information. Later approximately an hour was spent with each boy in administering the Wechsler-Bellevue Adult Scale. At some time during the remainder of the study, each boy was asked to complete the following six tests at a time and place of his own choosing: the SRA Youth Inventory, the Index of Adjustment and Values, Form A of the Kuder Preference Record (Personal), the High School Form of the Rotter Incomplete Sentences Blank, the Guilford-Zimmerman Temperament Survey, and the Thurstone Temperament Schedule.

Immediately following the administration of the "three-choice, three-criterion" sociometric test, all grade eleven students were asked to complete a copy of the Revised Ohio Social Acceptance Scale (Forlano and Wrightstone,³). On this scale the student was asked to assign each of his classmates a rating of from one to five depending on the degree to which they were socially attractive to him. From these data it was possible to compute the acceptance-by-group score and the acceptance-of-group score for each isolate and to compare these scores with the average scores for the respective isolates' classrooms.

An average of three and a quarter hours was spent in interviewing each of the fourteen subjects. During the early part of these interviews, the boys were encouraged to talk about themselves and their feelings, direct questions from the investigator being kept to a

minimum. As time went on, however, a more direct attempt was made to secure information about topics which did not seem to have been adequately discussed.

Parents and teachers were consulted for additional information. Teachers were asked to describe the isolates they taught, dealing with the isolate's achievement, his behavior, his relations with his fellow students, and his relations with his teacher. An average of nearly seven teachers described each isolate. Six homes were visited and the parents were invited to describe their child's development and personality. None of the parents was told the purpose of the study.

Subjects as Individuals

Sociometric Data

Identification of the social isolates chosen for this study was based entirely on data obtained by using the "three-choice, three-criterion" sociometric test mentioned above. Though the degree and nature of social isolation differed from subject to subject, all were below the fifteenth percentile of social acceptance on this scale. However, when sociometric status was measured by the Ohio Scale, a much greater variation appeared. The subjects ranged in acceptance by their classmates from the first to the sixtieth percentile. In view of the Ohio Scale results, there is some justification for questioning whether or not two of the subjects could really be classed as social isolates.

In addition, variation was present both in the awareness subjects had of their isolation and in their feelings about their isolation. Some confessed to a desire for more or better friends, while others expressed satisfaction with the existing situation. Some were painfully conscious of being socially isolated, while one insisted that he was very popular.

Considerable variation was also evident in the degree of acceptance the subjects accorded their classmates, some being highly accepting and others being very unaccepting of their classmates.

Uniqueness of Each Subject

The fourteen cases were compared with each other to determine whether or not those factors which might be expected to contribute to social isolation were unique in each case or common to several cases. A few unique factors were identified. One boy had unusually high intelligence which because of its rarity limited that boy's chances of finding persons interested in and capable of conversing at his level. An unusual and consuming interest set another boy apart from his fellows. A third boy was so involved in a round of part-time work, studies, and hobbies that he had almost no free

time and consequently little opportunity to cultivate intimate social relationships. It would probably be possible to extend this list of unique factors, but with little profit, since in a larger group of isolates, these factors might not be unique at all. It should also be noted that just because a factor seemed to be unique in one case, it was not necessarily the most significant factor in that case.

The search for unique factors led to two observations. First, although in some cases one subject appeared to be the only subject affected by a certain factor, there were usually other factors in his case which were common to other cases as well. In no instance did every one of the factors contributing to a subject's isolation appear to be operating in his case exclusively. Second, although the factors were not usually unique separately, in combination they did make each case different from the rest. In this respect, then, it is possible to say that each case was unique.

Subjects as a Group

An effort was made to determine what factors, if any, were common to the group of isolates as a whole. To this end the subjects were compared with respect to their scores on standardized tests, with respect to various easily observed aspects of their personality and environment, with respect to their feelings.

Standardized Tests

The percentile ranks of the lowest-scoring and highest-scoring isolates together with those of the second-lowest-scoring and second-highest-scoring isolates were compared on each trait measured by the standardized tests administered. (See Table I), The group seemed heterogeneous with respect to most traits, but seemed relatively homogeneous with respect to a few.

According to the results of the Guilford-Zimmerman Temperament Survey, with one or two exceptions, the isolates seemed to be relatively submissive, subjective, and unsociable or shy.

On the Thurstone Temperament Schedule, with the exception of one boy, the isolates appeared to be rather reflective.

The CRA Youth Inventory results showed that, with the odd exception, the subjects seemed to have a large number of personal problems, social problems, and health problems. All isolates scored above the fiftieth percentile on the Basic Difficulty Scale of this test, suggesting the possibility of serious personality difficulties.

Since a positive correlation between intelligence and social status has been reported by Thorpe (6), it is of interest to note that there seemed to be no lack of intelligence among the subjects of this study. Their full scale intelligence quotients on the Wechsler-Bellevue Adult Scale ranged from 103 to 136.

TABLE I
PERCENTILE RANKS OF THE LOWEST, HIGHEST, SECOND-
LOWEST, AND SECOND-HIGHEST-SCORING ISOLATES ON
THE SCALES OF FOUR STANDARDIZED TESTS

Test and Trait	Lowest Rank	Highest Rank	Second Lowest Rank	Second Highest Rank	Investigator's Judgement on the Homogeneity of Isolates With Respect to Each Trait
Guilford					
General Activity	1	79	2	73	heterogeneous
Restraint	3	91	5	70	heterogeneous
Ascendance	2	50	2	30	homogeneous
Sociability	1	64	2	28	homogeneous
Emotional Stability ..	1	70	2	70	heterogeneous
Objectivity	2	50	3	50	homogeneous
Friendliness	0	64	21	64	heterogeneous
Thoughtfulness	25	94	30	85	heterogeneous
Personal Relations ...	1	97	19	94	heterogeneous
Masculinity	14	91	17	72	heterogeneous
Thurstone					
Active	2	100	3	75	heterogeneous
Vigorous	1	97	4	86	heterogeneous
Impulsive	2	93	4	79	heterogeneous
Dominant	0	82	2	67	heterogeneous
Stable	1	90	16	76	heterogeneous
Sociable	1	89	2	65	heterogeneous
Reflective	48	99	58	99	homogeneous
Kuder					
Sociable	2	97	19	85	heterogeneous
Practical	2	76	7	73	heterogeneous
Theoretical	13	99	25	97	heterogeneous
Agreeable	5	84	5	77	heterogeneous
Dominant	5	97	7	86	heterogeneous
SRA					
School	5	96	12	80	heterogeneous
Future	3	80	12	78	heterogeneous
Personal	46	99	51	92	homogeneous
Social	28	99	46	88	homogeneous
Family	16	99	16	94	heterogeneous
Heterosexual	8	88	8	88	heterogeneous
Health	63	98	63	97	homogeneous
General	13	96	28	95	heterogeneous
Basic Difficulty	57	99	57	94	homogeneous

Environment, Background and School Achievement

Each isolate was considered with respect to about twenty fairly obvious aspects of his environment, background, and academic record. The two main points deserving attention are that eleven of

the fourteen boys had had poor school achievement in relation to their ability, and that eight had failed at least one subject taken in grade eleven. Seven of the fourteen isolates did not belong to any organized group of teen-age children either in or out of school, and six were members of one or another ethnic minority.

Feelings

A record was kept of the feelings expressed by the subjects either directly in speech or indirectly through their behavior and manner of expression. At least thirteen of the fourteen isolates seemed to feel insecure, and at least twelve lacked confidence either generally or in a particular area. Eleven boys felt inferior, and eleven felt some degree of loneliness and wanted friends. At least nine boys were anxious, and another nine were shy. Nine were dissatisfied with their social skills, such as their ability to converse or dance. Eight isolates admitted they were often unhappy.

Summary and Conclusions

The goal of this study was the identification of traits, circumstances, and other factors that were characteristic of a majority of the socially isolated subjects, and therefore possibly related in some way to social isolation. Characteristics which appeared to be common to a majority of the subjects included submissiveness, subjectivity, reflectiveness, unsociableness or shyness, and the possession of personal, social, and health problems. In addition, although three boys were earning high marks, most of the subjects were underachieving in school. However, the similarities in actual behavior were not as remarkable as the similarities in feelings and needs which motivated this behavior. Feelings of inadequacy, inferiority, and anxiety were mentioned frequently by the subjects. Needs for status, acceptance, and belonging were very much in evidence. Consequently, the observation that most subjects suffered from rather strong feelings of insecurity was inescapable. Hence, as far as adolescent boys are concerned, it seems possible that some relationship may exist between social isolation and feelings of insecurity. This conclusion can hardly be described as novel since it has been stated previously in one form or another by investigators working with younger subjects. Kuhlen and Bretsch (4), for example, reported as one of their findings that ninth-grade boys who were unaccepted evinced subjective social insecurity.

Although the goal of this study was the identification of factors common to the majority of the social isolates, it was recognized that there might be some factors related to social isolation which were not common to the majority. Some of these, such as unusual ability, unusual interests, and lack of leisure time, have already been

mentioned. But it should also be recognized that there might be some factors related to social isolation which did not operate in the case of any of the fourteen subjects selected for this particular study. For this reason, it must be understood that no claim is being made that the findings presented here constitute a complete list of all the factors related to social isolation.

At least sixty different factors have been linked with low social acceptance in the various studies that have dealt with this problem (Nielsen, 5). Although many of these factors appeared to be unrelated to the social isolation of the subjects in this study, practically all the observations made in this study were made and reported on in some previous study. However, none of the studies referred to dealt specifically with a group of senior high school boys, and none seemed to attach any great degree of significance to feelings of insecurity, so that in these respects the findings of this study may have some singularity.

Since the isolates in this study were not compared with more socially accepted students, it is not possible to state mathematically the degree of confidence that can be placed upon any of the observations made. Hence, these observations should be viewed as being no more than suggestive of what might be true. Possibly they may serve as a source of hypotheses for future experimental studies.

REFERENCES

1. Bloch, Herbert A. *Disorganization—Personal and Social*. New York: Alfred A. Knopf, 1962, 541.
2. Fleming, C. M. (ed.). *Studies in the Social Psychology of Adolescence*. London: Routledge and Kegan Paul Ltd., 1951, 144.
3. Forlano, G., and J. W. Wrightstone. "Measuring the Quality of Social Acceptability Within a Class." *Educational and Psychological Measurement*, 15, 1955, 127-136.
4. Kuhlen, Raymond G., and Howard S. Bretsch. "Sociometric Status and Personal Problems of Adolescents." *Sociometry*, 10, 1947, 122-132.
5. Nielsen, J. K. "Case Studies of Socially Isolated Males in Senior High School." Unpublished M.Ed. thesis, University of Alberta, 1962, 31-33.
6. Thorpe, J. G. "An Investigation into Some Correlates of Sociometric Status Within School Classes." *Sociometry*, 18, 1955, 49-61.

TEACHER CHARACTERISTICS AND GRADE IX SOCIAL STUDIES ACHIEVEMENT

WESLEY P. EDDY

The Problem

There appears to be general agreement that the position of the teacher is paramount in the teaching process and that quality of teachers is one of the most important factors in the success or failure of education.

The measurement of pupil achievement is a logical technique for measuring the productiveness of the teacher. Barr states that the change produced in pupils, measured in terms of the objectives of education, is the ultimate criterion of teaching success (2).

Changes produced in pupils may be variously defined in terms of different objectives of education. These changes may be attitudinal, or they may be results which can be objectively measured on a terminal examination.

At least one investigation indicated that the Alberta Department of Education Examinations in Grade IX are valid measures of educational achievement (4).

This study attempted to compare certain characteristics of the Grade IX social studies teachers in Alberta, classified according to various individual and situational factors, with the results of their students on the Social Studies IX final examination of 1958. The following characteristics were investigated in this study: (1) the extent of total academic and professional preparation beyond high school; (2) the number of university-level courses taken in the social studies; (3) years of teaching experience; (4) age; (5) the enrolment of the median-sized class taught; (6) the hours per week devoted to the teaching job, whether this be done at home, at school, or elsewhere; (7) the number of teachers in the school; (8) the sex of the teacher; (9) the teachers' subject-field preference for teaching.

The Main Problem. The main problem was whether any of the above characteristics of the Grade IX social studies teachers reflected any significant differences on the results of their pupils on the Grade IX final social studies examination.

The Subsidiary Problems. There were two subsidiary problems. (1) Were there any significant differences in the above characteristics between teachers in city schools and teachers in non-city schools? (2) Were there any significant differences of students' re-

sults in the Grade IX social studies examinations between students in city schools and students in non-city schools?

Background

This study parallels to some extent work done by S. A. Lindstedt and E. Wasylyk, both of whom used the same basic materials and investigated some of these same factors in the field of mathematics (11,16).

Review of Related Studies

Of the many teacher competence studies which have been made only a few used the product factor, or changes in terms of the student, as the criterion of measuring teacher competency. The investigations which have been made are sometimes contradictory. In 1945 Rostker conducted a study in which supervisors tested the pupils and rated twenty-eight seventh and eighth grade teachers in small Wisconsin schools on their pupils' gains in achievement on two units in citizenship. Knowledge of subject matter was found to be significantly associated with teaching ability (2). McCall and Krause sought to measure competence of seventy-three teachers against the growth of their sixth grade pupils in North Carolina. They found that the teachers' knowledge of subject matter showed no correlation with efficiency (13).

Stephens and Lichenstein, in their study of eighty-six grade five Baltimore arithmetic teachers, concluded that there was a slight correlation between class efficiency and years of teaching experience, especially for younger teachers (15). McCall and Krause found no such significant difference between years of teaching experience and pupil growth (13).

Lindstedt's Alberta study found no significant relationship between mathematics background and student achievement in Grade IX (11). Wasylyk, on the other hand, found a very significant increase in mathematics achievement with increase in the number of mathematics courses taken by their Alberta Grade XII teachers (16).

More conclusive findings have been made in some other areas. Where there are significant differences in the research findings, the following examples are typical. In a 1945 study by Brookover, sixty-six male Indiana high school teachers of United States history were studied. Pupil gains in information increased with the age of the teacher up to age thirty-eight after which they decreased (6).

Schunert, in his comprehensive study of 100 Minnesota secondary schools in 1951, reported that classes enrolling from twenty to thirty pupils exceeded the achievement of larger as well as smaller classes. He also found that classes in schools enrolling from 100 to

500 pupils exceeded the achievement of classes in larger and smaller schools (14). In 1945 Barr reported little, if any, evaluation of sex differences in teaching effectiveness had been made (3).

The need for further research is clearly evident.

Collection of Data and Method of Analysis of Data

A sample of schools was chosen at random. The Grade IX social studies teachers and students in these schools formed the samples used. The students had to be correctly identified according to their teachers either through the school records, or records of individual teachers verified by their students.

The Alberta Royal Commission on Education used a questionnaire to make a survey of the Alberta teacher force in 1958 (12). The questionnaire, which was returned by over ninety-nine per cent of the teachers, contained much factual information regarding the characteristics of the teachers and their schools. The information regarding the teachers which was used in this study was taken from the Alberta Royal Commission questionnaire.

The student marks, in stanine scores, in social studies were taken from the records of the Alberta Department of Education. This final mark assigned to the student by the Department of Education in 1958 was the average of the result of the final examination and the principal's confidential return mark in social studies. Several researchers have shown that the principals' marks are an accurate prediction of final examination marks and therefore the final assigned grading and the results on the final social studies examination are not appreciably different (5,10).

The scores used in this study are the scores of social studies achievement when mental ability was held constant. The mental ability test given by the High School Entrance Board of the Alberta Department of Education in 1958 was the School and College Ability Test (SCAT).

In all, ninety-nine teachers were included in the sample, or about 13.7 per cent of the total number of 710 Grade IX social studies teachers. Of these, nineteen or about 19.6 per cent reported that they were city teachers and 78 or about 80.4 per cent reported that they were non-city teachers.

There were a total of 2279 students in the sample, or about 13.1 per cent of the total of 17,378 Grade IX students in Alberta. Of these, 933 or about 40.9 per cent of the sample were city students and 1346 or about 59.1 per cent were non-city students.

Because only a sample of the total population of students was used, the total sample of students was tested for representativeness using the chi-square test of goodness of fit. The probability value

(.50 < p < .70) is greater than .05 indicating that the observed sample distribution was not significantly different from the expected distribution.

Findings

Analysis of variance reveals significant differences among the means of the groups indicating that there is a strong relationship between student results in social studies and the amount of training of the teacher. Using Tukey's gap test it can be shown specifically that there are no significant differences in social studies examination results between students of teachers with two or three (but less than four) years of academic and professional training and students whose teachers had one or less or six or more years of training. The achievement of students whose teachers had four or five but less than six years of training is very much higher than that of teachers with less training. The achievement of students whose teachers had six or more years of training is higher than the achievement of students whose teachers had one (less than two full years) or less years of training. The achievement of students whose teachers had six or more years of training is significantly lower than that of students whose teachers had four or five (but less than six) years of training.

There are no significant differences in Grade IX social studies examination results between students whose teachers had no university-level social studies courses and students whose teachers had one or two such courses. There is no significant difference between students whose teachers had three or four university-level social studies courses and students whose teachers had five or more such courses. The achievement of students whose teachers had five or more university-level social studies courses does not differ from the achievement of students whose teachers had no such courses. The achievement of students whose teachers had three or four social studies courses is very significantly higher than that of students whose teachers had no courses or one or two courses in social studies. The achievement of students whose teachers had five or more courses in social studies is very significantly higher than that of students whose teachers had one or two such courses.

Tukey's gap test showed very significant differences in social studies results between the students whose teachers had twenty-five or more years of teaching experience, and the students whose teachers had less than twenty-five years of experience. The results of the students whose teachers had twenty-five or more years of experience are higher than the results of the students whose teachers had less experience. There are no significant differences between students whose teachers had less experience.

The examination results are significantly higher in schools

where there are twenty-five or more teachers compared with schools with fewer teachers. The differences are not significant between schools of up to twenty-five teachers.

The Grade IX social studies final scores of students whose teachers preferred teaching in the social studies-English field are very much higher than those of students whose teachers preferred to teach in some other field.

No significant differences were found between Grade IX social studies marks of students and the age, median-sized class, hours per week devoted to the teaching job, and the sex of the teacher.

Table I shows the data for the two-way analysis of variance of the students' scores on the Grade IX social studies examination arranged according to the academic and professional training and the teaching experience of their teachers. The variance ratio for interaction is 2.56 which with 4 and 2259 degrees of freedom is significant at the .05 level.

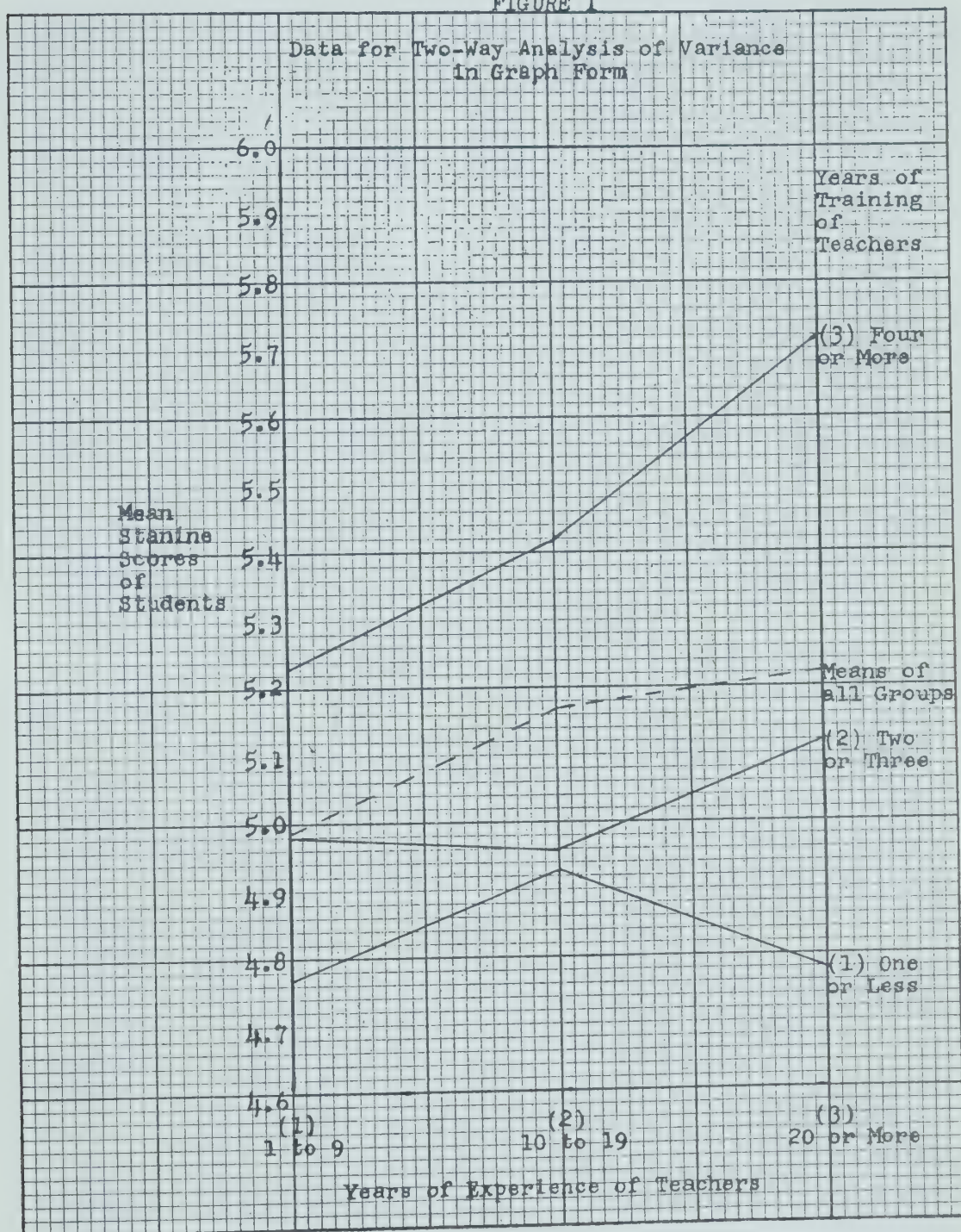
TABLE I
DATA FOR TWO-WAY ANALYSIS OF VARIANCE: GRADE IX
SOCIAL STUDIES RESULTS CLASSIFIED HORIZONTALLY
BY NUMBER OF YEARS TEACHING EXPERIENCE AND
VERTICALLY BY NUMBER OF YEARS OF ACADEMIC
AND PROFESSIONAL TRAINING OF TEACHERS

Years of Training		Years of Experience			Totals and Mean
		(1) 1 to 9	(2) 10 to 19	(3) 20 and over	
(1) One or less (less than two)	Number of Students ...	279	228	118	626
	Sum of Scores	1330	1130	564	3024
	Mean	4.77	4.93	4.78	4.83
	Variance	1.69	2.02	1.64	
(2) Two or three (less than four)	Number of Students ...	137	248	376	761
	Sum of Scores	682	1229	1927	3838
	Mean	4.98	4.96	5.12	5.04
	Variance	1.67	1.61	1.91	
(3) Four or more	Number of Students ...	273	427	181	881
	Sum of Scores	1429	2313	1035	4777
	Mean	5.23	5.42	5.72	5.42
	Variance	1.84	1.69	1.6	
Total					
Number of Students ...		689	904	675	2268
Sum of Scores		3441	4672	3526	11649
Mean		4.99	5.17	5.22	5.13

Homogeneity of Variance = Fmax = 1.26

Variance ratio for interaction = 2.56 F.01 = 3.32 F.05 = 2.37

FIGURE 1



This confirms what has already been established about the relationship between results of students and the amount of training of teachers and the number of years of experience of their teachers.

The variance ratio for interaction between these two factors indicates that there is a significant relation to the results of the students.

Figure I shows the data for two-way analysis of variance arranged in graph form. The mean results in social studies of students whose teachers had one or less (less than two full years) years of training increased from 4.77 to 4.93 as the teaching experience of teachers increased from one to nine to ten to nineteen years. The mean then decreased to 4.78 as the experience of the teachers increased to twenty or more years.

The mean results of students whose teachers had two or three (but less than four) years of training decreased from 4.98 to 4.96 as their teachers' experience increased from one to nine to ten to nineteen years. The mean increased to 5.12 as the experience of teachers increased to twenty or more years.

The mean results of the group of students whose teachers had four or more years of university training and one to nine years of

TABLE II
DIFFERENCES BETWEEN CITY SCHOOL TEACHERS AND
NON-CITY SCHOOL TEACHERS OF GRADE IX SOCIAL
STUDIES WITH RESPECT TO VARIOUS TEACHER
CHARACTERISTICS

Teacher Characteristic	Chi ²	df	p	Difference
Academic and Professional Training	12.92	3	.001 < p < .01	Very Sig.
Number of Social Studies Courses Taken	6.64	3	.05 < p < .10	Not Sig.
Number of Years of Teaching Experience	4.94	3	.10 < p < .20	Not Sig.
Age of the Teacher	6.02	3	.10 < p < .20	Not Sig.
Enrolment of the Median- Sized Class	26.3	3	.001 < p < .01	Very Sig.
Hours per Week Devoted to the Teaching Job	3.88	3	.20 < p < .30	Not Sig.
Number of Teachers in the School	29.57	3	.001 < p < .01	Very Sig.
Sex of the Teacher	.34	1	.50 < p < .70	Not Sig.
Subject-Field Preference of the Teacher	8.75	1	.001 < p < .01	Very Sig.
Differences with Respect to Exam- ination Results of Students	20.83	8	.001 < p < .0-	Very Sig.

experience was 5.23. This increased to 5.42 as the experience of the teachers increased to ten to nineteen years. As the experience increased to twenty or more years the mean rose to 5.72.

The mean of all the groups of scores of students increased from 4.99, to 5.17, to 5.22 as the experience of the teachers increased from one to nine, to ten to nineteen, to twenty or more years.

This shows that as experience increased the students' scores did not increase as rapidly for the whole population, if their teachers had less than four years of university training.

Table II shows the results of chi-square tests of independence. City teachers surpassed non-city teachers in number of years of academic and professional training, number of students in the median-sized class, number of teachers in the school, and subject-field preference. The two groups did not differ significantly in the number of university-level social studies courses taken, number of years of experience, age of the teachers, number of hours per week devoted to teaching, or sex of the teacher.

City students made significantly higher scores on the Grade IX social studies examination than non-city students.

Summary and Implications

The following were the more important findings and conclusions:

1. There was a significant positive relation between social studies marks made by pupils and the number of years of academic and professional training of, the number of university-level courses taken in the social studies by, the number of years of teaching experience of, and the subject-field preference of their teachers, and also the number of teachers in the school.
2. There was no significant relation between social studies marks of students and the age, median-sized class, hours per week devoted to the teaching job, and the sex of their teachers.
3. There was a significant interaction between amount of teacher training and length of teaching experience.
4. City teachers of social studies surpassed non-city teachers in number of years of academic and professional training, number of students in the median-sized class, number of teachers in the school, and subject-field preference for teaching in the English-social studies field.
5. There was not a significant difference between city teachers and non-city teachers in relation to the number of university-level social studies courses taken, number of years of teaching experience, age, number of hours devoted per week to teaching, and sex.

6. City students made significantly higher scores on the Grade IX social studies examination than non-city students.

The teacher characteristics which have a significant positive relationship with Grade IX student marks in social studies and are also significantly different between city and non-city teachers are the following: city teachers have more years of training and a greater preference for teaching in their subject field, and there are more teachers in an average city school compared to a non-city school.

It appears, then, that the reason city students obtain higher social studies marks than non-city students is that their teachers surpassed the non-city teachers in the above characteristics.

Downey reported in 1960 that Canadians (in Alberta), as a group assigned considerably higher priority than did Americans to such things as knowledge and scholarly attitudes (7). This may lead to rather different results in Canadian teacher effectiveness studies than in similar studies in the United States. There is, therefore, even more need for further research in the field of teacher effectiveness, particularly the effectiveness of Alberta teachers. Such studies could be carried out in other subject areas, and at various grade levels as well as using behavioral factors.

The results of this study imply that it is desirable for Grade IX social studies teachers to have more years of academic and professional training, more university-level courses taken in the social studies, more years of teaching experience, and a preference for teaching in the English-social studies field. The results also imply that it is desirable for teachers of Grade IX social studies to have four or more years of training, including at least three content courses in the social studies.

This study also seems to imply that the important step in training is at the four year or degree level.

While many of the teachers reported in this study no doubt taught in their subject preference area as well as other areas, most teachers teaching social studies preferred to teach in some other area. Specifically 25 per cent in city schools and 65.3 per cent in non-city schools for an astonishing total of 58.2 per cent of the teachers in the survey had a teaching preference in some subject-field other than English-social studies. Yet teachers appear to do the most effective teaching in the subject-field in which they prefer to teach and so should be assigned to that field.

REFERENCES

1. *Annual Report, 1958*. Edmonton: Department of Education, Government of Alberta, The Queen's Printer, 1959.
2. Barr, A. S., and others. *The Measurement of Teaching Ability*. Madison, Wisconsin. Denbar Publications, Incorporated, 1945.

3. Barr, A. S., and others. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations," *Journal of Experimental Education* XXX (September, 1961), 5-156.
4. Black, D. B. "The Prediction of University Freshman Success Using Grade IX Departmental Examination Scores," *Alberta Journal of Educational Research*, V (December, 1959), 229-239.
5., "A Study of the Relationship of the Grade IX Principal's Rating to Performance on the Alberta Grade IX Departmental Examinations," *Alberta Journal of Educational Research*, IV (December, 1958), 227-236.
6. Brookover, W. B. "The Relation of Social Factors to Teaching Ability," *Journal of Experimental Education*, XIII (1954), 191-205.
7. Downey, L. W. "Regional Variations in Educational Viewpoint," *Alberta Journal of Educational Research*, VI (December, 1960), 195-199.
8. Eddy, W. P. "A Study of Certain Characteristics of Teachers in Relation to Grade IX Social Studies Achievement." Unpublished Master's thesis, The University of Alberta, Edmonton, 1962.
9. Edwards, A. L. *Statistical Methods in the Behavioral Sciences*. New York: Rinehart and Company, Incorporated, 1954.
10. Evenson, A. B. "Re: Grade IX Examinations." Curriculum Branch. Edmonton: 1952. (Mimeographed).
11. Lindstedt, S.A. "Teacher Qualifications and Grade IX Mathematics Achievement." Unpublished Master's thesis, The University of Alberta, Edmonton, 1960.
12. MacArthur, R. S., and S. A. Lindstedt. *The Alberta Teacher Force in 1957-58*. University of Alberta Monographs in Education Number 3. Edmonton: The Alberta Advisory Committee on Educational Research, University of Alberta, 1958.
13. McCall, W. A., and G. R. Krause, "Measurement of Teacher Merit for Salary Purposes," *Journal of Educational Research* LIII (October, 1959), 73-75.
14. Schunert, J. "The Association of Mathematical Achievement with Certain Factors Resident in the Teacher, in the Teaching, in the Pupil, and in the School," *Journal of Experimental Education*, XIX (March, 1951), 219-238.
15. Stephens, J. M., and A. Lichenstein. "Factors Associated with Success in Teaching Grade Five Arithmetic," *Journal of Educational Research* XL (May, 1947), 683-694.
16. Wasylyk, E. "The Relation Between Four Selected Teacher Characteristics and Student Achievement in Grade Twelve Mathematics." Unpublished Master's thesis, The University of Alberta, Edmonton, 1961.

PHYSICAL FITNESS OF MENTALLY RETARDED BOYS

WAYNE L. SENGSTOCK

University of North Carolina

The purpose of this investigation was to see if educable mentally retarded boys differ from intellectually normal boys of comparable chronological age, and intellectually normal boys of comparable mental age in the performance of tests of physical fitness.

There are two reasons which make it important to investigate the many factors present in motor performance and their relationship to intelligence. First, knowledge is needed about how mentally retarded children compare with normal children in this area if the best possible programs of physical education are to be made available for these children. Second, there is a need for more positive knowledge concerning the relationship between intelligence and motor performance.

Physical development and motor performance affect an individual's economic, social, and emotional adjustment. Knowledge of a particular group's deficiency in the area of physical development and/or motor performance precedes the development of programs to reduce the deficiency. The ability of the mentally retarded to perform gross motor tasks must, therefore, be investigated before more effective physical education programs can be developed. Good programs of physical education can help the mentally retarded become more capable in the performance of tasks requiring physical fitness. Adequate physical fitness, therefore, increases the potential of the mentally retarded for making a satisfactory economic, social, and emotional adjustment.

Rabin (7) reported that motor skills have important implications for vocational placement. This is true for the mentally retarded in particular. "Employment of an intellectual nature, is closed to him, whereas he has some accessibility to the more routine mechanical jobs in which motor skills oft-times play a crucial role."

There are many jobs which require a degree of intelligence which the mentally retarded do not possess. For instance, one could not expect the mentally retarded to participate in those jobs classified as highly skilled or those occupations requiring college training. Many jobs, however, are available which require only a limited degree of intelligence and the performance of certain motor skills which rely on physical fitness. Jobs which qualify under this description can be found in heavy construction industries, the building trades, food service and related areas.

In a study of adolescent motor performance Espenschade (3) stated: "Contacts with others before the adolescent years are frequently made through active play. Thus the child makes friends, becomes part of the group. Excellence in physical performance may win approval of the group and of the parents and of other adults." Jones and Bayley (5), from a study relating physical maturing among boys to behavior, concluded:

Those who are physically accelerated are usually accepted and treated by adults and other children as more mature. They appear to have relatively little need to strive for status. From their ranks come the outstanding student body leaders in high school. In contrast, the physically retarded boys exhibit many forms of relatively immature behavior: this may in part be because others tend to treat them as the little boys they appear to be. Furthermore a fair proportion of these boys give evidence of needing to counteract their physical disadvantage in some way—usually by greater activity and striving for attention although in some cases by withdrawal.

Examples of how an individual's motor performance can affect his economic, social, and emotional adjustment have been described by Rabin, Espenschade, and Jones and Bayley. An investigation of how the mentally retarded children compare with intellectually normal children must precede the development of the best possible programs of physical education for mentally retarded children.

Another reason for this investigation was the conflicting attitudes expressed by writers concerning the relationship between intelligence and motor performance. One group of authors found that there is little relationship between the two, while others have reported a positive correlation between intelligence and the ability to perform well on motor tasks.

Kirk and Johnson (6) recognized the lack of agreement concerning the relationship of motor performance and motor skill to intelligence. Being concerned with identification and diagnosis of mental retardation insofar as it pertained to educational planning for this group of schoolage children, they stated: "In the field of mental retardation additional information concerning the relationship of intelligence and motor ability is especially desired."

Sherman (10) stated:

Defective children can learn simple sensory-motor activities as quickly and efficiently as bright children, and with proper training retarded children will be as adequate as normal children in performance requiring gross sensorimotor adjustment.

In the area of industrial psychology, Tiffin (12) supported this view in a reference to mechanical aptitude requirements of a job: He said:

The layman assumes that the person who is "clever with his head" is also clever with his hands" and that the individual who is high in mental ability is likely to be high in body dexterity, muscular coordination, and

aptitude for learning skilled movements. Such is not the case. It is impossible to predict whether an individual who is high in one will be high in the other.

On the other hand, some authors have claimed that the relationship between motor performance and intelligence is more than casual. Bolton (2) stated, "In-so-far as an individual is wanting in motor development, he is wanting in mental development."

Tregold (13) further supported this hypothesis about the mentally retarded when he stated:

Probably, however, the most frequent and marked abnormality of movement in defectives is inco-ordination. The co-ordination of movement aims at securing the desired end and with a maximum of precision and a minimum of effort. It is dependent upon a proper development of both sensory and motor pathways; but even where this exists as in the normal child, perfect co-ordination is only attained by constant practice. In persons suffering from even the mildest degree of amentia co-ordination is often acquired with difficulty, and remains imperfect; and although many of them learn to use their hands with considerable amount of dexterity, the balance and movement of the body often continue clumsy and ungainly.

Kirk and Johnson have pointed out the need for research to determine the relationship between intelligence and performance on motor tasks. Some writers, like Sherman and Tiffin, have expressed the opinion that there is little relationship between intelligence and motor performance. Other writers, like Bolton and Tredgold, have supported the hypothesis that a high correlation exists between these two variables.

These examples from the literature help to identify two reasons why the investigation of motor performance of the retarded is important. The literature indicates a need for research in this area so that realistic educational programs can be developed. Secondly, more positive knowledge concerning the relationship between intelligence and motor performance is needed.

The present study is concerned with the area of physical fitness, which is related to motor performance. Measuring the performance of mentally retarded boys and comparing it with the performance of intellectually normal boys on tests of physical fitness can help to increase our knowledge in the above mentioned areas.

Hypotheses

There were two general hypotheses proposed in this study.

1. The physical fitness of an educable mentally retarded group of boys is significantly different from the physical fitness of a group of intellectually normal boys of comparable chronological age.
2. The physical fitness of an educable mentally retarded group of boys is not significantly different from the physical fitness of a group of intellectually normal boys of comparable mental age.

These hypotheses were tested by comparing the means of the various groups on each of 7 tests of a physical fitness test battery with the analysis of variance. Critical differences were set at the .05 level. For these data an F - ratio of 4.18 is significant at the .05 level, and an F - ratio of 7.60 is significant at the .01 level.

Related Research

In reviewing the research in the field no studies were found comparing the physical fitness of mentally retarded children with the physical fitness of intellectually normal children. Most of the early studies relating motor performance to intelligence were involved with the measurement of reaction time, and simple motor response or dexterity. Several of the techniques used were highly subjective, which casts some doubts on the validity of the data. Many of these early studies did not use mentally retarded subjects and were performed with intellectually normal and superior subjects.

When the retarded were finally studied for motor performance it was primarily reaction time which was investigated. Gradually other facets of motor performance were studied, such as "motor skills," "motor development," "motor proficiency," "muscular skills," and "motor characteristics." Although no studies were found which measured physical fitness, per se, many of the test items employed in the above studies are similar to those used in tests of physical fitness. These studies give valuable insights into the motor performance of the mentally retarded, but none of them investigated all the components of physical fitness. Of all types of motor performance, physical fitness is one of the most important in terms of practical application.

The studies which are most noteworthy are those of Beaber (1), Francis and Rarick (8), Thurstone (11), and Howe (4). Beaber compared the performance of a group of mentally retarded children and the performance of two groups of normal children on four tests of "simple fine motor ability." The results showed that the performance of the retarded group was below that of normal children of similar C.A. and was similar to the performance of a group of intellectually normal children of comparable M.A. The studies of Francis and Rarick, Thurstone, and Howe compared the gross motor performance of mentally retarded children with the performance of normal children of comparable C.A.'s. In each of these studies the performance of the mentally retarded was significantly below the performance of intellectually normal children.

In many programs of special education for the educable mentally retarded, the children take physical education classes with normal children of the same age under the assumption that in motor per-

formance the retarded are similar to normal children of comparable chronological age. This study will attempt to investigate this assumption.

Procedure

Subjects of the Study

Three groups of thirty boys were selected for this study. These groups were designed as Educable Mentally Retarded (EMR), Old Normal (ON), and, Young Normal (YN) groups. The EMR and the ON groups were equated on the variable of chronological age, and the EMR and YN groups were equated on the variable of mental age. The three groups of subjects used in this study were drawn from public school districts in Onondaga County, New York. All subjects selected were boys who had no record of, or observable, uncorrected physical or sensory impairment that would affect performance on any of the tests. Only those subjects completing all of the tests were used in the study.

All subjects of the EMR group were enrolled in special classes for the mentally retarded at the time of testing. These classes were located in the same buildings as were the regular classes and were considered a part of the total school program.

The EMR boys met the additional criteria of a C.A. within the range of 10 years 0 months through 15 years 0 months, and WISC full scale I.Q. within the range of 60-80 at the time the battery of motor tests was administered.

All special classes for educable mentally retarded children that contained boys within the prescribed age range were considered in selection of the EMR group. From the total, a group of schools was selected at random until the required number of educable mentally retarded boys was available. Five schools were selected before this criterion was met.

The ON group of subjects was matched with the EMR group of subjects on the basis of CA and had a WISC full scale I.Q. within the range of 90 - 110 at the time the battery of motor tests was administered.

All of the ON boys in the same schools as an EMR boy, having a CA within two months of the EMR subject were considered. From this group, subjects were selected at random and administered the WISC until a boy was found who had an I.Q. within the range required. There were four instances where ON boys meeting these criteria were not available. In each of these cases there was no normal boy in the school old enough to be matched with an EMR boy. In these four cases, ON boys were selected who had previously attended the school involved, but were now in attendance at ad-

vanced schools in the district. The procedure of random selection, as described above was used for these four boys also.

The YN boys were matched with EMR boys on the basis of MA and had WISC full scale I.Q.'s within the range of 90 - 110 at the time the battery of motor tests was administered.

All of the YN boys in the same school as an EMR boy, having a CA similar to the MA of the EMR child, and within the I.Q. range required by this study, were considered. From this group, subjects were selected at random and administered the WISC until a boy was found having a mental age within two months of the EMR boy.

Data for the group means, standard deviations, and ranges on CA, MA, and I.Q. are contained in Table I.

TABLE I
STATUS OF THE THREE GROUPS ON CHRONOLOGICAL AGE, MENTAL AGE, AND INTELLIGENCE QUOTIENT

Group	Chronological Age (Months)			Mental Age (Months)			Intelligence Quotient (Full Scale)		
	$\bar{X}$	σ	Range	$\bar{X}$	σ	Range	$\bar{X}$	σ	Range
YN	111.47	14.29	83-139	114.13	12.77	87-146	102.73	5.14	91-110
EMR	158.83	15.25	131-184	114.10	12.63	89-147	72.13	5.96	60- 80
ON	158.60	15.25	129-182	157.80	15.23	120-190	99.53	4.58	92-107

The families of each of the subjects were evaluated on the basis of socio-economic class with Warner's "Index of Status Characteristics" (14). The socio-economic class ratings were analyzed against the X^2 distribution to test the hypothesis that there were significant differences between groups. With 4 degrees of freedom $X^2 .95 = 9.49$, and $X^2 .90 = 7.78$. The observed X^2 value of 3.57 was therefore non-significant at these confidence levels and it was concluded that there were no significant differences between groups on the basis of socio-economic class.

Test Conditions and Procedures

The America Association of Health, Physical Education and Recreation Youth Fitness Test Battery was selected as the instrument to measure physical fitness in this study for several reasons. It is accepted as a valid measuring instrument by the most prominent leaders in the field. This test measures speed, strength, coordination, and endurance. It was selected for its ease in administration and scoring and because the individual tests require a minimum of instruction and are easily comprehensible by the subjects.

The individual tests of the battery are:

- 1. Pull-Ups
- 2. Sit-Ups
- 3. Shuttle-Run
- 4. Standing Broad Jump
- 5. 50-Yard Dash
- 6. Softball Throw for Distance
- 7. 600-Yard Run-Walk

The physical fitness tests were administered on the playgrounds of the various schools attended by the subjects. Equipment used were stop watches and a tape measure. These tests were administered during physical education classes. The complete battery of *AAHPER Youth Fitness Test* was administered to each child on successive days, with a rest period of 3 to 7 minutes after each test. The Sit-Ups, and Pull-Ups, Standing Broad Jump, and Softball Throw were administered individually. The Shuttle-Run, 50-Yard Dash, and 600-Yard Run-Walk were administered in pairs. Each test was explained and demonstrated to the subjects prior to administration. The subjects copied the demonstration to allow the examiner to correct faulty performance and to minimize the differences in comprehension between subjects. The subjects were told that they were competing against time rather than against their partners.

Results

In comparing the performance of the EMR group with the performance of the two intellectually normal groups, raw scores were used. Table 2 summarizes the performances of the three groups.

TABLE II
RAW SCORE GROUP MEANS ON THE AAHPER TESTS

Test	Raw Score Term	Group		
		YN	EMR	ON
Pull-Ups	Number of Repetitions	1.60	1.73	2.87
Sit-Ups	Number of Repetitions	37.83	37.13	58.53
Shuttle Run	Time in seconds	12.79	12.16	11.11
Standing Broad Jump	Distance in inches	89.80	57.23	67.27
50-Yard Dash	Time in seconds	9.15	8.38	7.87
Softball Throw	Distance in feet	70.33	99.13	130.07
600-Yard Run-Walk	Time in minutes and seconds	2-56	2-41	2-22

Analysis of the raw score data indicates that the performance of the ON group was superior to the performance of the EMR group in each of the seven tests. In comparing the performances of the EMR and YN groups, the performances of the EMR group was superior to that of the YN group of six of the seven tests.

The differences between groups were tested with the analysis of variances. Table III shows the results of the analysis of variance of the group performances on the tests of physical fitness.

TABLE III
F - RATIOS BETWEEN GROUPS ON THE AAHPER
YOUTH FITNESS TESTS

Test	EMR - YN	EMR - ON
Pull-Ups	.06	4.05
Sit-Ups	.01	17.05**
Shuttle-Run	7.89**	14.81**
Standing Broad Jump	17.07**	12.20**
50-Yard Dash	14.87**	7.09*
Softball Throw	23.73**	12.25**
600-Yard Run-Walk	14.52**	20.52**

d.05 = 4.18

d.01 = 7.60

Statistical analysis of the raw scores failed to reject hypothesis 1 in all but one of the tests, the Pull-Ups. A possible explanation for obtaining no significant differences on the Pull-Up Test can be found in the nature of the test. Because of the difficulty encountered by the subjects in performing Pull-Ups, the mean scores were very low, the range very small, and an exceptional performance by one or two boys was able to disrupt the homogeneity of variance. Despite the lack of statistical significance, the ON group was able to perform on the average one pull-up more per boy than the EMR group.

The findings of the present study rejected hypothesis 2 in five of seven tests. That is to say the performances of the EMR group was significantly superior to the performances of the YN group in all but the Pull-Ups and Sit-Ups tests.

Conclusions and Discussion

This investigation shows that in all instances the ON group of subjects was superior to the EMR group in performing tests of physical fitness. On six of the seven tests this superiority was

statistically significant. The performance of the EMR group was superior to the performance of the YN group on five of the seven physical fitness tests. Therefore, on the basis of this investigation, educable mentally retarded boys, of the age groups measured, differed significantly from intellectually normal boys of comparable chronological age, and from intellectually normal boys of comparable mental age in the performance of the physical fitness tests used in this study. Over-all the EMR group performance was inferior to the ON group and superior to the YN group.

The results of this study have important implications for physical education programming of mentally retarded children. This investigation has identified certain groups of mentally retarded children as being inferior in physical fitness as measured by the test used in this study. Under certain conditions the retarded are slower, weaker, less coordinated, and have less endurance than normal children of the same age. Therefore, the implication is that the policy of placing EMR boys with their CA peers for physical education is open to serious question if the assumption underlying the program is that the two groups' motor performance should be comparable.

The results of this study indicate that there is a relationship between intelligence and motor performance, but to what extent cannot be determined. The EMR group was superior to the YN group with which they were matched, but the overwhelming advantages of height and weight of the EMR group contaminated this aspect of the study.

Before broader interpretation can be accepted, subsequent studies utilizing more extensive populations, and measures of other types of motor performance must be made. Investigations of how the mentally retarded compare with intellectually normal children in motor ability, motor skill, or motor development would be of assistance in making more accurate judgments about performances of EMR groups.

It is suggested in future studies that subjects from resident training schools as well as subjects from a variety of cultural areas be included. Comparisons should also be made using more diverse socio-economic levels and other chronological age groups. Other studies using matched groups equated on the basis of height and weight are also recommended.

There are many questions raised by the present study. To answer these questions, additional studies are suggested. Some of the questions raised are given below. If the mentally retarded are, generally, physically unfit, what instrumental arrangement for physical education training would best meet the needs of the re-

tarded? Would it be more beneficial to segregate the mentally retarded, or to integrate them with normal children for physical education? If integration is suggested, with which age group should they attend classes? Should they attend classes with normal children of their own age or with younger normal children? If programs for the mentally retarded are to reach maximum potential in the area of physical performance, the above questions must someday be answered.

REFERENCES

1. Beaber, James, "The Performance of Educable Mentally Handicapped and Intellectually Normal Children on Selected Tasks Involving Single Motor Performance" (Unpublished Ed.D. dissertation, School of Education, Syracuse University, 1960.)
2. Bolton, T. L. "The Relation of Motor Power to Intelligence," *American Journal of Psychology*, XIV (October, 1903), pp. 366-367.
3. Espenschade, Anna "Motor Performance in Adolescence," *Psychological Studies of Human Development*, ed. Raymond G. Kuhlen and George C. Thompson (New York: Appleton-Century-Crofts, 1952), pp. 55-56.
4. Howe, Clifford E. "Comparison of Motor Skills of Mentally Retarded Children and Normal Children," *Exceptional Children* XXV (1959), 352-54.
5. Jones, Mary S., and Bayley, Nancy. "Physical Maturing Among Boys as Related to Behavior," *Psychological Studies of Human Development* ed. R. G. Kuhlen and G. C. Thompson (New York: Appleton-Century-Crofts, 1952), p. 48.
6. Kirk, Samuel, and Johnson, Orville G. *Educating the Retarded Child* (New York: Houghton Mifflin Co., 1951), p. 55.
7. Rabin, Herbert H. "The Relationship of Age, Intelligence, and Sex to Motor Proficiency in Mental Defectives," *American Journal of Mental Deficiency*, LXII (November, 1957) p. 507.
8. Robert J. Francis and Rarick, Lawrence S. *Motor Characteristics of the Mentally Retarded*. Cooperative Research Monograph No. 1, U.S. Department of Health, Education, and Welfare (Washington U.S. Government Printing Office, 1960.)
9. Sengstock, Wayne L. "A Comparison of the Performance of Educable Mentally Retarded Boys with the Performance of Intellectually Normal Boys on the American Association for Health, Physical Education, and Recreation Youth Fitness Test Battery" (Unpublished Ed.D. dissertation, School of Education, Syracuse University, 1963.)
10. Sherman, Mandel *Intelligence and Its Deviations* (New York: Ronald Press, 1945), p. 24.
11. Thurstone, Thelma G. *An Evaluation of Educating Mentally Handicapped Children in Special Classes and in Regular Classes* U.S. Office of Education (Washington: U.S. Government Printing Office, 1959.)
12. Tiffin, Joseph *Industrial Psychology* (2nd ed. rev.; New York: Prentice-Hall, Inc., 1947), p. 122.
13. Tredgold, A. F. *A Textbook of Mental Deficiency*, (New York: Houghton Mifflin Co., 1951), p. 55.
14. Warner, W. Lloyd; Meeker, Marchia; and Eels Kenneth. *Social Class in America* (New York: Harper and Brothers 1949.)

ESSAY GRADING AS A FUNCTION OF WRITTEN LANGUAGE SKILL

ARTHUR G. STOREY
University of Alberta, Calgary

Rationale and Hypotheses

During July, 1963 the present author was charged with the task of teaching 22 Division II and III Manitoba teachers "Utility and Creative Writing" for 19, fifty minute periods. The perennial problems besetting all teachers of language accompanied the task—i.e. how to measure the gains made in written language in any valid way? Of course a teacher of language can, and unfortunately often does, avoid the whole issue of both measurement and writing by teaching grammar, for one can measure what a student knows *about* language with comparative ease while the assessment of what he can do *with* language remains elusive. (Thorndike and Hagen (1), like many other authors in the field of measurement and language give a comprehensive account of the difficulties inherent in the essay as a measurement instrument).

Having just completed the marking of some 660 pupil themes in connection with the W. J. Gage, Calgary Study before undertaking the "Utility and Creative Writing" task, the present author was keenly aware of the difficulties concerned with language evaluation. This awareness together with the availability of the large number of student themes written under standardized conditions instigated the hypothesis that: (1) Before taking the "Utility and Creative Writing" course the Manitoba teachers enrolled in it would grade essays too high, i.e. higher than the present author graded them when comparing them with models based on those provided for the *Cooperative Sequential Tests of Educational Progress*, Essay Form 3. (2) After completing the writing course the teachers would grade comparative essays (a) significantly lower than they had before taking the course, and (b) with a mean not significantly different from that awarded by the author.

Procedure

Before the first meeting of the writing class the 660 Grade VII and VIII essays written on two parallel topics were shuffled together for purposes of randomization and placed in 22 folders, each containing 30 essays. These folders were then assigned to the 22 students making up the writing class. The contents of each folder was then sub-divided for the purpose of "before" and "after" the "Utility and Creative Writing" course marking.

At the first meeting of the class each student was given a folder containing 15 of the essays assigned to him. The class was then informed that there would be certain assignments made during the course and that the first one would be the grading of the essays they had been given. The marking instructions were:

I want you to grade these papers as fairly as you can. Do not consider mechanics of language but only the skill with which the ideas are presented. Pretend the essays were written by your own class and mark them as fairly and as consistently as you can. Read all of them first and sort them into A, B, C, D, and E piles then go through them again and award them marks in multiples of five. Please return them to me at the beginning of tomorrow's class.

On the following day the essays were collected and nothing more was said about theme marking until the 18th meeting of the class. At this time each student was given the second half of the papers assigned to him and was asked to mark them. The above instructions were again read to the class. Table I summarizes the grades awarded both before and after the writing course.

TABLE I
RESULTS OF BEFORE AND AFTER MARKING

Score Interval	Frequency Before Course	Frequency After Course
93-100	5	0
85- 92	34	6
77- 84	32	14
69- 76	84	48
61- 68	37	39
53- 60	69	77
45- 52	41	53
37- 44	15	21
29- 36	6	24
21- 28	4	16
13- 20	0	18
5- 12	3	14
Mean	64.72	51.55
S.D.	15.94	19.14
t 9.82 (P < .01)		

An examination of Table I indicates that the Manitoba teachers awarded the 330 essays marked before the writing course a mean of 64.72 which is 18.40 higher than the author's mean of 56.32. Since the t of this difference is 6.94 the hypothesis that the teachers would mark too high is confirmed—the assumption being that the author's mean of 56.32 approximates the true value of the essays.

The teachers' "after" the writing course mean of 51.55 is 13.17 points lower than their "before" mean. This difference yields a t of 9.82 which confirms part (a) of the second hypothesis, i.e. the teachers "after" mark is significantly lower than their "before" mark.

The difference between the teachers' "after" mean and the author's mean is 4.77; the teachers' mark being the lower of the two. This difference yields a t of 4.68 making it necessary to reject the null hypothesis that there would be no difference between the teachers' "after" mark and that awarded by the author.

The curriculum of the "Utility and Creative Writing Course" stressed vocabulary usage, sentencing, and the organization of ideas. The students were given as much practice in these skills as possible through paragraph and essay writing. Table II gives the gains made

TABLE II
GAIN IN WRITING SKILL AND DECREMENT IN MARKS
AWARDED TO PUPIL ESSAYS

Student	Skill Increase	Mark Decrement
1	15	18
2	10	13
3	10	12
4	15	14
5	5	3
6	20	35
7	20	22
8	5	0
9	10	4
10	5	8
11	15	15
12	15	12
13	25	21
14	10	13
15	15	10
16	10	8
17	10	13
18	20	11
19	10	11
20	15	17
21	20	12
22	25	16
Mean	13.87	13.09

by each student between his first to final essay. It also gives the decrement in the marks awarded by the students to the Grade VII and VIII essays discussed earlier.

The rank order coefficient of correlation between the increase in language usage skill and the decrement in marks awarded pupil themes is .68 indicating a marked relationship ($P < .01$) between the two. Both means are significantly different from zero at the .01 level. There is no significant difference between the means.

Conclusions

The present study demonstrates, then, that validity in essay marking is at least in some degree a function of the marker's own skill in the use of written language.

BIBLIOGRAPHY

1. Thorndike, R. L. and Hagen E. *Measurement and Evaluation in Psychology and Education* (2nd. ed.) John Wiley and Sons, New York, 1961, pp. 41.
2. Adkins, Wood, Dorothy, *Test Construction Development and Interpretation of Achievement Tests*, Charles E. Merrill Books, Inc., Columbus, Ohio, 1961 Passim.
3. Gerberich, J. R. et al., *Measurement and Evaluation in the Modern School*, David McKay, Inc., New York 1962 Passim.
4. Mirrielees, L. B., *Teaching Composition on Literature*, (Revised ed.) Harcourt, Brace and Co., New York 1952, Ch 1.
5. Dawson, A. D. and Zollinger, M. *Guiding Language Learning*, Harcourt, Brace and Co., New York, 1957, Passim.

CONTENTS

Volume IX, 1963

No. 1, March

	PAGE
A National Testing Program for University Admission	3
<i>Donald B. Black</i>	
Case Studies of High School Drop-Outs	13
<i>Gordon J. Rancier</i>	
Parental Behavior and Student Attitudes Toward Further Training	22
<i>Henry Zentner</i>	
Use of Films in Divisional and County Schools in Alberta	31
<i>George Kravetz and H. T. Sparby</i>	
The Safran Culture Reduced Intelligence Test	36
<i>C. Safran</i>	
Changes in Attitude and Behavior as a Function of Methods Courses	45
<i>Arthur G. Storey</i>	
Performance in Arithmetic of Edmonton, Alberta and Christchurch, New Zealand Schools	49
<i>Donald B. Black and Marion E. Foster</i>	
Book Reviews	60
<i>J. W. Gilles</i>	
<i>W. B. Dockrell</i>	

No. 2, June

	PAGE
Changes in Patterns of Thinking in Elementary Arithmetic	65
<i>S. A. Lindstedt</i>	
Selection of Children for Remedial Teaching of Reading	73
<i>Barry P. Frost</i>	
Cultural Assimilation Between Indians and Non-Indians	79
<i>Henry Zentner</i>	
Hobbes and Modern Humanism	87
<i>Ian Housego</i>	
Selecting Administrators	94
<i>John E. Davis</i>	
Efficiency of Predictors at the University of Alberta	100
<i>Laurence L. Mack</i>	
An Applied Linguistic Method in Grade Seven French	111
<i>Sister Simon-Hermann, A.S.V.</i>	
Book Reviews	120
<i>M. A. Nay</i>	
<i>J. H. M. Andrews</i>	

No. 3, September

	PAGE
The Role of Stimulus Affect on Concept Attainment	125
<i>David R. Olson and R. J. C. Harper</i>	
Some Personality Characteristics of Education Students	132
<i>Barry P. Frost</i>	
The Construction and Use of a Guidance Attitude Scale	140
<i>C. Safran</i>	
Two Hypotheses Concerning Normalization of Form Perception	147
<i>Harvey W. Zingle and Roland A. Lambert</i>	
Social Class and Ninth Grade Educational Achievement	157
<i>Thomas E. Linton and Donald F. Swift</i>	
Value Congruence Among Indian and Non-Indian High School Students	168
<i>Henry Zentner</i>	
Leader Behavior of Principals, Staff Morale and Productivity	179
<i>B. T. Keeler and J. H. M. Andrews</i>	

No. 4, December

	PAGE
"Unified Phonics Method" in Teaching Reading in Grade I	195
<i>A. Fred Deverell</i>	
Free-Writing in Grades IV, V and VI	225
<i>Florence I. McKie and M. D. Jenkinson</i>	
Some Factors and Achievement in Literature 20	238
<i>John Hawrelko</i>	
Case Studies of Socially Isolated Males	247
<i>J. K. Nielsen</i>	
Teacher Characteristics and Social Studies Achievement	254
<i>Wesley P. Eddy</i>	
Physical Fitness of Mentally Retarded Boys	264
<i>Wayne L. Sengstock</i>	
Essay Grading as a Function of Written Language Skill	274
<i>Arthur G. Storey</i>	

The Editorial Committee for the
ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

will be pleased to consider research articles from contributors outside the Province of Alberta, for publication in the Journal.

Manuscripts should be addressed to the Editor.

H. E. Smith

